ARQUEOLOGÍA COMPUTACIONAL DEL TERRITORIO

Métodos y técnicas para estudiar decisiones humanas en paisajes pretéritos

Miguel Carrero-Pazos



ARQUEOLOGÍA COMPUTACIONAL DEL TERRITORIO

Métodos y técnicas para estudiar decisiones humanas en paisajes pretéritos

Miguel Carrero-Pazos

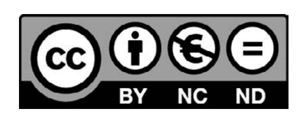
Archaeopress Archaeology



Archaeopress Publishing Ltd
Summertown Pavilion
18-24 Middle Way
Summertown
Oxford OX2 7LG
www.archaeopress.com

ISBN 978-1-80327-632-8
ISBN 978-1-80327-633-5 (e-Pdf)

Cover: Diseño de la portada por Diego Rodríguez Robredo.



A Andrew Bevan y Mark Lake,
Enrico Crema, Quentin Bourgeois
y Carlos Rellán

Este trabajo ha sido realizado gracias al plan postdoctoral I2C – Xunta de Galicia (Universidade de Santiago de Compostela) y el programa de investigación e innovación Horizonte 2020 de la Unión Europea, acuerdo de subvención nº 886793, Marie Skłodowska-Curie IF-EF-ST 2019, *MegaScapes* (*University College London*)





Índice

Presentación

La Arqueología se encuentra hoy plenamente inserta en la era informática y digital. Este proceso, común al conjunto de las Ciencias, está suponiendo una importante renovación metodológica de la disciplina, hecho que se materializa en el surgimiento de nuevas líneas de trabajo que auguran futuros prometedores.

En los últimos años, la ciencia arqueológica viene haciendo frente a un auténtico aluvión de datos llegados de proyectos de investigación presentes y pasados: es el Big Data. Los métodos de computación juegan, en este panorama, un papel clave. La existencia de proyectos que abarcan períodos históricos con cronologías amplias, regiones supranacionales o incluso porciones enteras del globo es, hoy, la normalidad. Instituciones políticas actuales, como la Unión Europea, financian cada vez más este tipo de investigaciones ambiciosas. Se trata del famoso slogan *high risk, high gain*: el migracionismo en la Prehistoria, la extinción de los neandertales y la llegada del ser humano anatómicamente moderno, la propagación del cultivo del arroz en Asia o la expansión del fenómeno megalítico europeo son, entre muchos, proyectos que utilizan una perspectiva "top-down", analizando miles de datos de regiones amplias del planeta con métodos y técnicas diversas.

Proyectos como los citados son ejemplos también de cómo la interdisciplinariedad es, en la actualidad, una característica básica y necesaria de todo equipo de investigación. Estamos en un momento crucial para la arqueogenética, la bioarqueología, la arqueología satelital o la arqueología computacional, disciplina, esta última, sobre la que versa este libro. Ninguna de las citadas son campos de trabajo nuevos, sino que cuentan con una dilatada historia. Es patente, sin embargo, que hoy todas disfrutan de un proceso de renovación que está permitiendo la reformulación de preguntas históricas pasadas, junto al planteamiento de otras nuevas que antes eran inabarcables o impensables.

La hiperespecialización del investigador también es señal de que la disciplina arqueológica continúa evolucionando. En la actualidad es frecuente encontrar, por ejemplo, a arqueólogos e historiadores de formación especializados en la aplicación de técnicas informáticas sin importar el período histórico. La situación a la inversa también es común, con informáticos o físicos de formación que desarrollan su carrera por entero en la Arqueología. Todo ello es síntoma de la madurez que muestra nuestra disciplina, que proporciona a los investigadores múltiples caminos a seguir a la vez que la enriquecen.

Por otro lado, y como parte de una tendencia generalizada en la Ciencia moderna, la Arqueología está sumándose, aunque quizás a paso lento, a un proceso de transparencia y apertura en toda su estructura. Las investigaciones que suelen calificarse como "opacas", en las que no se comparten los datos arqueológicos utilizados o no se describen con detalle los métodos de análisis, están cediendo terreno ante aquellas investigaciones que sí lo hacen. Se trata de hacer que el conocimiento científico sea fácilmente localizable, accesible, interoperable y reutilizable, principios que definen la iniciativa FAIR.[1] De hecho, el papel de instituciones de investigación supranacionales, como la *European Research Commission* (ERC), en promocionar una investigación pública, en abierto, y transparente, está siendo clave para esta transformación. No poco se está haciendo en este sentido, pues cualquier proyecto de investigación financiado por este organismo debe tener un plan de publicación *open access*, con lo que son quizás las viejas mentalidades de los investigadores, a veces proteccionistas, las que deben cambiar. Una ciencia social como es la Arqueología debe promover, ante todo, el acceso público a la información, tanto la primaria (datos en bruto procedentes de excavaciones, prospecciones, etc.), como de los métodos y técnicas utilizados en las investigaciones y publicaciones.

Con este telón de fondo, la Arqueología computacional, también llamada *Arqueo-informática* es, sin lugar a duda, una de las subdisciplinas arqueológicas con mayor desarrollo en los últimos años, a la par de la revolución informática y digital.

Este libro ha sido diseñado, ante todo, para el investigador y estudiante de Historia y Arqueología que busque iniciarse o complementar su formación con la aplicación de métodos y técnicas computacionales, en concreto aquellas relacionadas con el estudio y evolución de territorios y paisajes culturales pretéritos. Se suma, de esta forma, a una larga lista de monografías y artículos de investigación que evidencian que las técnicas informáticas y estadísticas son una parte fundamental en la ciencia arqueológica actual.

El texto que sigue confiere un protagonismo esencial a las explicaciones metodológicas. Estas son el principal objetivo y ocupan la mayor parte de un discurso que se ha simplificado al máximo, al tiempo que no se olvida el planteamiento de casuísticas arqueológicas específicas.

La estructura del libro responde a una lectura desde lo teórico hacia lo práctico, y lo simple a lo complejo. De esta forma, el texto se torna ejemplo, en forma y contenido, de cómo encarar un estudio en Arqueología del territorio y paisaje con herramientas computacionales, con dos casos prácticos que pueden ser utilizados y adaptados fácilmente por los estudiantes e investigadores en sus casuísticas particulares. Dicha estructura es muestra, también, de la evolución del autor como arqueólogo especializado en el uso de estas herramientas.

Formado en la Universidad de Santiago de Compostela, en el seno del grupo de estudios megalíticos dirigido por el profesor Antón A. Rodríguez Casal, participó en campañas de prospección arqueológica como parte del equipo de campo, en las que pudo aprender los métodos tradicionales de registro como la cartografía en papel, el uso del GPS o la fotografía aérea más tarde, hasta el análisis de los resultados de campo con herramientas de teledetección.

[1] https://www.go-fair.org/go-fair-initiative/

Técnicas que le han permitido conocer, de primera mano y en profundidad, los entresijos de una actividad arqueológica fundamental, como es la prospección.

Desde aquí, el uso de los Sistemas de Información Geográfica para contrastar hipótesis sobre la localización de los megalitos gallegos, tema de su tesis doctoral, le ha llevado, en los últimos años, a centrarse en la modelización y simulación estadística de los paisajes megalíticos atlánticos, a través del uso de la programación estadística con *R Statistics*.

La primera parte del libro está dedicada a la adquisición de datos y aspectos teóricos. El capítulo 1 expone los principales avances en prospección arqueológica de los últimos años. Se ofrece una breve puesta al día de esta subdisciplina, que nos lleva a considerar los últimos trabajos en teledetección o las visualizaciones derivadas del uso de la tecnología LiDAR.

El capítulo 2 aborda el marco teórico de la Arqueología del territorio y paisaje a través del uso de los Sistemas de Información Geográfica. Se evidencia que los SIG han sido herramientas de gran relevancia para la Arqueología desde sus primeras aplicaciones, no sólo desde el punto de vista metodológico sino también epistemológico. Esta evolución ha sido tal que un camino pedregoso, con importantes debates en torno a los límites y posibilidades del conocimiento generado con estas herramientas, que llegan incluso hasta la más reciente actualidad.

Una de las líneas de investigación donde los SIG han supuesto un auténtico punto de inflexión ha sido la Arqueología espacial y, siendo específicos, lo que respecta al *locational analysis*: estudiar las razones, criterios y decisiones tomadas por las sociedades pretéritas a la hora de construir, por ejemplo, sus asentamientos o monumentos funerarios. Estas cuestiones se abordan en el capítulo 3, haciendo hincapié en que las modernas técnicas estadísticas de simulación están permitiendo la renovación de los métodos hacia nuevos caminos analíticos que, sin lugar a duda, permiten salvar algunos de los problemas y entresijos técnicos más clásicos. Nuevas vías de trabajo que parecen discurrir hacia el uso de la programación estadística en conjunción con los SIG de código libre.

La segunda parte del libro se concreta en explicaciones metodológicas para el análisis y modelización de variables geográficas y culturales con Sistemas de Información Geográfica, su caracterización mediante estadística multivariante y modelado predictivo, y su estudio con análisis de patrones de puntos. Así, los capítulos 4 y 5 proporcionan claves para llevar a cabo estudios sobre la localización de sitios arqueológicos a partir de modelos interpretativos desarrollados por la historiografía, su factorización con herramientas SIG y su análisis mediante modelado predictivo y patrones espaciales de puntos.

La tercera y última parte está dedicada a la ciencia en abierto, transparente y reproducible, temas de suma importancia en la actualidad. El capítulo 6 proporciona explicaciones generales orientadas a entender el por qué debemos realizar investigaciones arqueológicas en abierto, huyendo de mentalidades proteccionistas que no hacen otra cosa que frenar el desarrollo y expansión del conocimiento, al tiempo que reducen su democratización.

En el capítulo 7 se expone un caso práctico en el que se estudian ciertos patrones de localización de los conjuntos megalíticos de As Chans de Barbanza (Galicia, Noroeste de la Península Ibérica), ejemplificando cómo puede realizarse investigación transparente y reproducible

mediante el uso de *R Statistics* y *Rmarkdown.* Se adjuntan, a través de un repositorio con acceso público, todos los datos utilizados en dicha investigación.

El capítulo 8 se ocupa de un segundo caso práctico. Se analiza el poblamiento castreño de la Edad del Hierro del Noroeste peninsular a través del estudio de una selección de castros localizados en tres comarcas gallegas. Se propone el análisis de ciertas variables físicas y orográficas junto al uso de la simulación de patrones de puntos para aproximarnos a la posible territorialidad existente entre estos asentamientos.

El capítulo 9 proporciona unas reflexiones finales acerca de la Arqueología computacional del territorio y paisaje, el uso de software libre e investigaciones reproducibles. Se trata de dejar patente la necesidad de crear trabajos transparentes que puedan ser utilizados, adaptados o replicados por cualquiera. Sólo con ello podremos generar un auténtico conocimiento científico compartido, que no entienda de fronteras administrativas, económicas, lobbies de mercado o el mero azar.

Apostemos por una ciencia arqueológica transparente, en abierto y reproducible, una Arqueología para el siglo XXI.

Agradecimientos

Este libro ha pasado por muchos estados desde que comenzó a ser pergeñado hace ya unos cuantos años. En ese proceso de corregir errores, rehacer capítulos y reorientar partes diversas jugó un papel fundamental mi carrera postdoctoral. Las bases están, sin embargo, en mi tesis doctoral y como tal debe agradecimiento a los profesores Antón A. Rodríguez Casal y Marcos Martinón-Torres. El primero supo orientarme desde mis tiempos de estudiante y el segundo me abrió un futuro académico acogiéndome en uno de los departamentos de arqueología más importantes del mundo, como es el Instituto de Arqueología de la *University College London*, del que, con privilegio y en años posteriores, pude formar parte en calidad de investigador del programa postdoctoral europeo *Marie Skłodowska-Curie*.

En la UCL tuve la ocasión de trabajar con los profesores Andrew Bevan y Mark Lake, referentes en Arqueología computacional. Quienes conozcan sus trabajos notarán su evidente influencia en las páginas que siguen. Las anotaciones que pude tomar de sus clases del *MSc Computational Archaeology: GIS, Data Science and Complexity*, uno de los másteres más punteros en el campo de la Arqueología computacional, son, en parte, el germen de varios capítulos de este libro.

Por otra parte, un contrato postdoctoral del gobierno autonómico de la Xunta de Galicia (Plan I2C) en la Universidade de Santiago de Compostela, me brindó la oportunidad de trabajar durante varios años en el extranjero, en concreto, con Quentin Bourgeois en la Universidad de Leiden (Países Bajos) y con Enrico Crema en la Universidad de Cambridge (Reino Unido). Ambas estancias de investigación fueron claves para mi formación porque me permitieron acceder a conocimientos avanzados tanto en Sistemas de Información Geográfica como en simulación estadística con *R*. Realizar el curso *From Data to Interpretation*, impartido por el segundo en el *McDonald Institute for Archaeological Research* de la Universidad de Cambridge supuso un verdadero punto de inflexión, ya que me permitió acceder al interesante campo de la estadística bayesiana.

Mención importante también para los profesores Germán Delibes de Castro, Ramón Fábregas Valcarce y Javier Fernández Eraso, cuyas sugerencias, apuntes y objeciones a mi texto doctoral supusieron una importante masa crítica para el trabajo que sigue.

Finalmente, me gustaría dejar constancia del agradecimiento a quienes desde un principio dedicaron tiempo a mi formación, bien a través de cursos, charlas, bien a través de discusiones sobre la potencialidad de las herramientas informáticas en el campo arqueológico. Muchas

de esas conversaciones han supuesto la reformulación, más de una vez, de alguno de los capítulos que siguen. Me refiero, sin importar el orden, a Emiliana Romaní, Carlos Díaz, Carlos Rellán, Benito Vilas, Alia Vázquez, Miguel Busto, David Espinosa, Fernando Rodríguez del Cueto, Marco de la Rasilla, Francesc Conesa, Hèctor Orengo, Elías López Romero, Francisco Cebreiro, Mikel Díaz, César Parcero y Pastor Fábrega. Por último, también a Clara Benítez, que vio madurar este libro muy de cerca.

En Santiago de Compostela, primavera de 2023

Miguel Carrero-Pazos

Universidade de Santiago de Compostela

PRIMERA PARTE:

TEORÍA Y MÉTODOS EN ARQUEOLOGÍA COMPUTACIONAL DEL TERRITORIO Y PAISAJE

Capítulo 1.

Prospección arqueológica remota y la “revolución” del LiDAR

La prospección arqueológica puede definirse como el conjunto de métodos orientados al reconocimiento de evidencias arqueológicas en superficie e inmediatamente por debajo. Esta labor puede realizarse a múltiples escalas y con diferente intensidad, con trabajos orientados, por ejemplo, al reconocimiento de estructuras arqueológicas en un territorio extenso, un paraje o paisaje determinado (denominada prospección extensiva), o aquellos que tratan de descubrir objetos materiales como restos cerámicos, en superficies de territorio más reducidas (prospección intensiva, o de cobertura total).

Los métodos utilizados abarcan toda una serie de ciencias anexas a la Arqueología, desde técnicas de teledetección e imagen digital (imágenes multiespectrales, tecnología LiDAR), hasta georradares o radiodetectores, que permiten constatar la existencia de estructuras bajo tierra sin necesidad de excavar. En la actualidad se están desarrollando aplicaciones de detección automática de fragmentos cerámicos en el terreno a partir de algoritmos basados en imágenes de alta resolución, obtenidas con drones (Orengo, García-Molsosa 2019), lo que supone un importante avance en la adquisición de datos arqueológicos en prospecciones intensivas.

Desde un punto de vista histórico, la prospección arqueológica superficial continuaba, hasta no hace mucho, siendo entendida como un paso previo a la excavación (Ruiz Zapatero 1988: 35; Ruiz Zapatero, Burillo Mozota 1988: 47; Fernández Martínez 1989: 54; García Sanjuán 2004: 186), y ello es así porque desde sus inicios como disciplina científica en el S. XIX, la Arqueología venía centrándose fundamentalmente en la excavación. Como indica E. Cerrato Casado (2011: 152), la prospección representaba una tarea menor, más propia de aficionados que de historiadores profesionales o académicos (Ruiz Zapatero 1988: 34-35).

Sin embargo, la aparición de la Arqueología procesual en la década de 1960 coadyuvará el desarrollo teórico y metodológico de este campo de investigación, que en España llegaría a partir de los ochenta con las contribuciones de académicos como G. Ruiz Zapatero (1988), F. Burillo Mozota (1988) o V. Fernández Martínez (1989), entre otros. Sus trabajos sentaron las bases de una disciplina que, a partir de los años noventa y con el cambio de siglo, sufriría un

proceso de renovación metodológica a la par que las tecnologías informáticas comenzaban un despegue imparable. Pronto se puso de relieve cómo la prospección, lejos de ser un conjunto de métodos científicos que pueden ser aplicados con mayor o menor rigor en la búsqueda de yacimientos arqueológicos, se constituía como un campo de investigación emancipado dentro de la disciplina arqueológica, contando con un fondo debate teórico y metodológico que contribuía en muy mucho al planteamiento y resolución de preguntas histórico-arqueológicas. Tal es el resultado, por ejemplo, de la combinación de estrategias metodológicas de la prospección arqueológica con un marco epistemológico emanado de la Arqueología espacial, posibilitando el estudio de las relaciones que se establecen en los espacios geográficos, entre diferentes evidencias arqueológicas, y entre éstas y el medio físico en el que se encuentran (Burillo Mozota 2013: 17).

1.1. El cambio de siglo. Estrategias adaptadas a una renovación tecnológica

Previo a la introducción del GPS, la mayor parte de los trabajos de prospección arqueológica se ocupaban de cartografiar sitios y elementos arqueológicos en base a técnicas manuales, como descripciones más o menos formales de la localización de los yacimientos o mediante técnicas cartográficas simples, como trazar rumbos. La prospección superficial se planificaba en torno a fases de gabinete, de revisión cartográfica, bibliográfica, y creación de bases de datos, junto a trabajo de campo orientado a la comprobación de las referencias recogidas. Como ejemplo podemos citar los proyectos del grupo de estudios megalíticos de la Universidade de Santiago de Compostela, liderado por el profesor A. A. Rodríguez Casal, que estructuraban la prospección arqueológica en base a los siguientes estadios analíticos (Rodríguez Casal *et al.* 1997: 527):

- Análisis y vaciado bibliográfico.
- Estudio y vaciado del catálogo patrimonial procedente de la administración pública (Servicio de Patrimonio).
- Estudio y documentación de la toponimia.
- Estudio de la documentación cartográfica: información geográfica, ambiental, topográfica, geológica y edafológica.
- Estudio de la capacidad productiva de los suelos.
- Estudio de la distribución de cultivos y aprovechamientos actuales.

Este proyecto de investigación, realizado a finales de la década de 1990 y que nos vale como ejemplo, se caracterizaba por desarrollar un tipo de prospección arqueológica superficial de corte tradicional, con trabajos que seguían la tendencia general del momento. De hecho, estamos ante el modelo de prospección arqueológica imperante en muchos territorios, y continúa siendo uno de los más utilizados en la actualidad, como puede observarse en los últimos trabajos que incluyen prospección arqueológica relativa, por ejemplo, al Megalitismo gallego, algunos derivados del sector profesional (por ejemplo, López-Felpeto Gómez 2014), y otros centrados en la investigación (Ferreira de Sousa 2013; Carrero Pazos *et al.* 2014).

Por otro lado, este modelo de prospección superficial no ha sido el único llevado a cabo. La Arqueología del paisaje en Galicia, personificada en el grupo de investigación dirigido por el profesor F. Criado Boado, planteó un discurso metodológico renovado que concluía que la prospección arqueológica superficial, "antes de buscar yacimientos, debe mirar el paisaje"

(Criado Boado *et al.* 1991: 49). Para ello, definieron una prospección intensiva y sistemática que se caracterizaba por “batir toda la superficie del territorio seleccionado para intentar descubrir yacimientos de todas las épocas” (Criado Boado *et al.* 1991: 58). Definían, además, una prospección que denominaron intensiva basándose en dos ejes fundamentales: la realización de encuestas muy minuciosas a los vecinos de la zona y, en segundo lugar, la inspección del terreno centrada en las particularidades del paisaje, con especial interés en los puntos con remoción de tierras para inspeccionar el subsuelo (Criado Boado *et al.* 1988: 244).

Otros trabajos posteriores han diseñado una confluencia entre ambas metodologías. Así, X. Ayán Vila y X. Amado Reino (2001) combinan la prospección superficial extensiva catalogando los yacimientos arqueológicos de un área de estudio determinada para evaluar el estado de conservación de los que sean visibles, con una prospección intensiva orientada al descubrimiento de restos arqueológicos de todos los períodos, realizando estimaciones sobre zonas potenciales de aparición de yacimientos que no son detectables a simple vista. Esta combinación daría lugar al planteamiento de una tercera opción metodológica de los sistemas de prospección superficial, en este caso, denominada prospección intensiva selectiva, caracterizada, según los autores (Méndez Fernández *et al.* 1995: 294; Ayán Vila, Amado Reino 2001: 130), por ocuparse de la identificación del registro arqueológico visible (prospección extensiva), y la localización de una muestra significativa de yacimientos “invisibles” (prospección intensiva en ciertas zonas).

Lo cierto es que todas las metodologías de prospección descritas son igual de válidas dependiendo de los objetivos de la investigación. Las primeras son quizás más adecuadas para el estudio de los yacimientos megalíticos desde un punto de vista macro, la segunda (y sus derivaciones), para la prospección en zonas más localizadas o relativas a proyectos más concretos que quieran orientar sus trabajos a estudios sobre cualquier yacimiento, de cualquier época, que pueda aparecer.

De lo que no cabe duda es que realizar una prospección arqueológica intensiva, con sistemas de prospección propios de zonas áridas (Criado Boado *et al.* 1988: 245), es muy difícil en territorios tan boscosos como el Noroeste de la Península Ibérica. Habida cuenta de que Galicia es uno de los territorios con mayor espesor vegetativo y menor porcentaje de suelo cultivado (entre un 20-35% del total) (Criado Boado *et al.* 1991: 58), es metodológicamente improbable conseguir realizar una “lectura” total del territorio. Y esto es así porque sencillamente nos encontramos con una cobertura vegetal tan espesa que provoca que, muchas veces, estructuras positivas como túmulos megalíticos de 1,5 metros de altura sean inidentificables durante el trabajo de campo.

1.2. El despegue de la tecnología: la teledetección y cartografía digital

En Arqueología, la teledetección puede definirse como el estudio y análisis de imágenes satelitales con el objetivo de buscar restos de estructuras arqueológicas superficiales, a diferentes escalas (Wilkinson 2003; Zubrow 2007).

El desarrollo de la informática con el surgimiento de Internet y la *World Wide Web* en los años noventa del siglo pasado marcó el inicio de una revolución en los estudios cartográficos que, de forma paulatina, iban incorporando nuevas herramientas a su *modus operandi*. El uso del

GPS para cartografiar yacimientos arqueológicos fue siendo introducido, hasta el punto de convertirse en la herramienta más importante en los trabajos de prospección. Todo ello acompañado de la aparición y masificación de las herramientas de teledetección, como las fotografías satelitales, abiertas al público en general gracias a programas como *Google Earth*, lanzado en su primera versión en 2001.

Desde aquí, los proyectos arqueológicos se extendieron con la creación de las denominadas plataformas *webmapping* o visores web de cartografía, que bebían de imágenes proporcionadas por diferentes satélites en cada país. Información cartográfica accesible y disponible de forma gratuita gracias a una interfaz amigable, como es el caso, por ejemplo, de *Google Earth*.

La disciplina arqueológica también se benefició de este tipo de plataformas. El uso de las fotografías aéreas satelitales, desde globos aerostáticos, dirigibles, cometas o más recientemente drones, para la detección remota de yacimientos arqueológicos y creación de planimetrías en detalle, cuenta con una tradición dilatada en el tiempo.[1] España, por su parte, no se ha quedado atrás este tipo de avances. Contamos tanto con investigaciones teóricas (Sánchez Palencia, Orejas Saco del Valle 1991; Orejas Saco del Valle 1995; Fernández García 2000) como con proyectos más específicos, que destacaron por desarrollar métodos de trabajo derivados de la prospección aérea e interpretación fotográfica.[2]

En lo que respecta a las imágenes satelitales, sólo con el desarrollo de la informática en las últimas décadas su uso se ha extendido y democratizado. Así, Google Earth es, sin lugar a duda, una de las plataformas de imagen satelital más utilizadas en Arqueología, a juzgar por su presencia en las publicaciones (Parcak 2009). Otra que ha jugado un papel importante en nuestro país es el “Sistema Gráfico Digital de Identificación de Parcelas Agrícolas” (SIGPAC) que, estructurado como un SIG, utiliza ortoimágenes aéreas (Figura 1).

El uso y democratización de las imágenes satelitales y aéreas fue uno de los avances más importantes de este tipo de visores *webmapping*, porque permitía gestionar, a la vez, planos digitalizados, información raster y capas vectoriales. Las imágenes que esta plataforma y otras utilizan, como IBERPIX,[3] pertenecen al Plan Nacional de Ortofotografía Aérea (PNOA),[4] cofinanciado y cooperado entre la Administración General del Estado (AGE) y las diferentes comunidades autónomas, a su vez enmarcado en el Plan Nacional de Observación del Territorio (PNOT)[5] coordinado por el Instituto Geográfico Nacional (IGN)[6] y el Centro Nacional de Información Geográfica (CNIG).[7]

Gracias a la directiva *Inspire* para el establecimiento de una infraestructura de datos geográficos en Europa, el desarrollo del PNOA supuso un avance fundamental para las ciencias geográficas de nuestro país, con objetivos como la obtención de ortofotografías aéreas digitales de alta

1 Al respecto pueden consultarse, entre otros, Deuel 1973; St. Joseph 1966; Anderson 1979; Wilson 2000; Bourgeois, Marc 2003; Brophy, Cowley 2005; Verhoeven, Loenders 2006.
2 Cfr. Rosselló Bordoy 1970; Sánchez Palencia, Fernández Posse 1992; Peña Monne *et al.* 1986; Martínez del Pozo *et al.* 2013; Costa García 2015; Carrero-Pazos, Vilas-Estévez 2016, entre otros.
3 https://www.ign.es/iberpix/visor/
4 https://pnoa.ign.es/
5 https://www.ign.es/web/plan-nacional-de-observacion-del-territorio
6 https://www.ign.es/web/ign/portal
7 https://www.ign.es/web/ign/portal/qsm-cnig

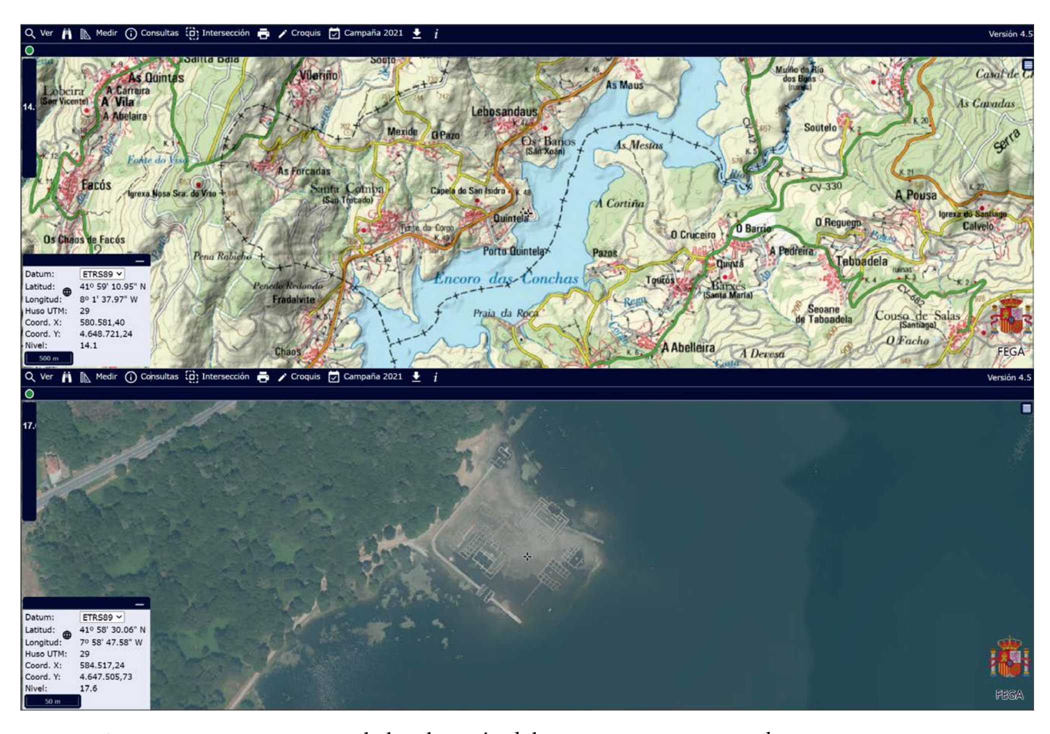


Figura 1. Visor SIGPAC con la localización del campamento romano de *Aquis Querquennis*, Bande (Ourense).

precisión para todo el territorio español, manteniendo un período de actualización de 2-3 años dependiendo de la zona (**Figura 2**).

La producción cartográfica se realiza de manera descentralizada por cada equipo autonómico, y se divide en dos grandes grupos: la generación de ortoimágenes y la realización de un vuelo fotogramétrico. El vuelo fotogramétrico equivale a una escala de 1:15000 (PNOA25) y 1:30000 (PNOA50), utilizando una cámara de alta resolución equipada con sensor pancromático y cuatro sensores multiespectrales (**Figura 3**). Además, se proporcionan los datos obtenidos con el sensor LiDAR.

Por último, es necesario llamar la atención sobre el PNOA Histórico, gestionado por el Centro Cartográfico y Fotográfico del Ejército del Aire (CECAF),[8] donde se ofrecen las imágenes aéreas de los vuelos fotogramétricos realizados por el territorio español desde el año 1930 (ortofotografía histórica).[9] En la actualidad, existen toda una serie de vuelos disponibles escaneados a alta resolución, aunque sin lugar a dudas, vuelos de suma importancia para la disciplina arqueológica son los denominados Americano Serie A (1945-1946) y Americano Serie B (1956-1957), puestos a disposición y consulta pública a través de varias plataformas *webmapping*, como son la Fototeca Digital[10] o, por ejemplo en el caso de Galicia, la última

[8] https://ejercitodelaire.defensa.gob.es/EA/ejercitodelaire/es/organizacion/unidades/unidad/Centro-Cartografico-y-Fotografico-CECAF/

[9] https://pnoa.ign.es/web/portal/pnoa-imagen/vuelos-y-ortofotos-historicas

[10] https://www.ign.es/web/sc-cnig-fototeca

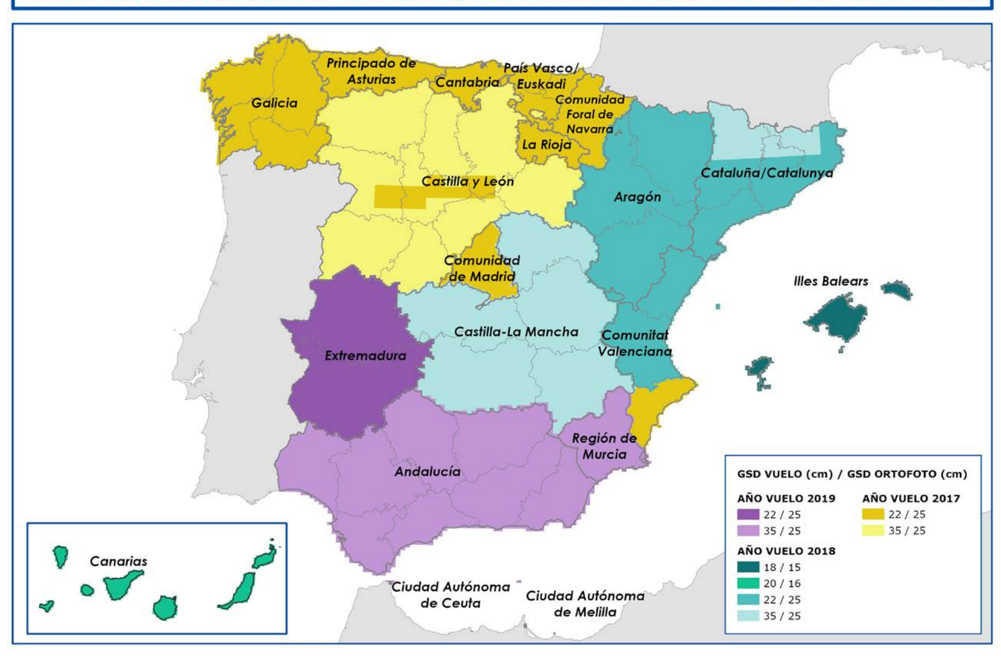


Figura 2. Proyecto PNOA. Año de vuelo de la última ortofotografía validada (Copyright proyecto PNOA. Ministerio de Transportes, Movilidad y Agenda Urbana, Gobierno de España).

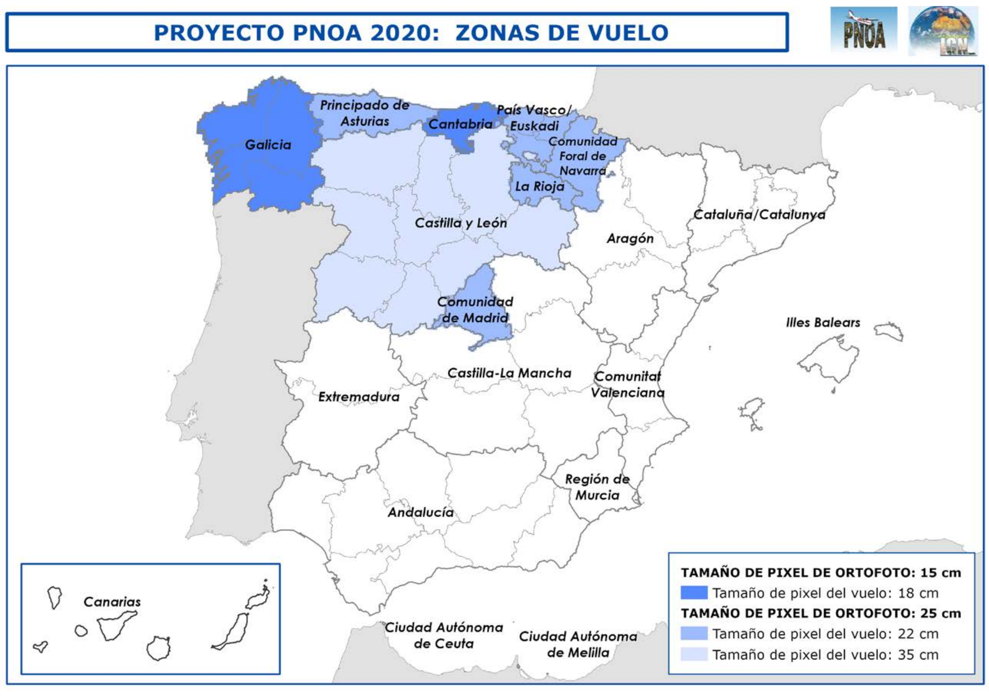


Figura 3. Zonas del vuelo fotogramétrico del PNOA para el año 2020 (Copyright proyecto PNOA. Ministerio de Transportes, Movilidad y Agenda Urbana, gobierno de España).



Figura 4. Comparativa del Castro de Santa Trega (A Guarda, Pontevedra) con el visor *webmapping* de la Xunta de Galicia. **Izquierda:** fotografía aérea del Vuelo Americano de 1956. **Derecha:** imagen del PNOA, de 2017.

aplicación cartográfica actualizada por el gobierno autonómico de la Xunta y el Instituto de Estudos do Territorio.[11] En ella, aparte de poder comparar las imágenes aéreas del Vuelo Americano Serie B (1956-1957) con ortoimágenes actuales de forma automática[12] (**Figura 4**), podemos analizar la superficie del terreno gallego gracias al uso de un Modelo Digital de Elevaciones sin vegetación y con sombreado analítico de gran resolución (5m), procedente del tratamiento de datos LiDAR.

El uso de la fotografía histórica en Arqueología ha sido recurrente en los últimos años, al permitir la reconstrucción de la topografía del terreno previa a los cambios urbanísticos

[11] http://mapas.xunta.es/visores/basico/
[12] http://mapas.xunta.es/visores/comparador/

contemporáneos. Su uso es fundamental a la hora de estudiar yacimientos arqueológicos desaparecidos, ya que, de esta forma, pueden llegar a localizarse e incluso reconstruirse. Es el caso, por ejemplo, del valle del río Salas (Baixa Limia, Ourense), en el que la construcción de un embalse en los años setenta anegó uno de los conjuntos megalíticos más interesantes de la zona (Eguileta Franco 1999; Ferreira de Sousa 2013). Gracias al uso de técnicas 3D y la fotografía histórica, se pudo detectar y cartografiar, con precisión, el conjunto de monumentos (Vilas Estévez *et al.* 2015).

El uso de la fotografía aérea también supuso un cambio importante en la planificación de la prospección arqueológica. Hasta ese momento se utilizaba la cartografía en papel para moverse por el campo, sin embargo, a medida que las plataformas *webmapping* se fueron extendiendo, el uso de ortoimágenes extraídas del SIGPAC, por ejemplo, fue imponiéndose como una buena forma de planificar los trabajos y acceder más rápido a los yacimientos, puesto que permitía, entre otras cosas, observar las carreteras y caminos de acceso, elementos que son cruciales a la hora de planificar una salida de campo efectiva. No obstante, con este tipo de visores no se podía ir más allá y obtener información de la superficie del suelo (*ground-truthing*, en la terminología inglesa más utilizada), dado el espesor del manto vegetativo. En la actualidad, los trabajos que se aproximan a la detección de yacimientos mediante el estudio de imágenes satelitales, multiespectrales, de alta resolución tomadas con drones, junto con estudios de detección de sitios arqueológicos basados en datos LiDAR, ponen de relieve que la prospección arqueológica está disfrutando, hoy, de un importante *momentum* tanto teórico como metodológico.[13]

1.3. La tecnología LiDAR y su uso en Arqueología

En su devenir histórico, las metodologías de prospección arqueológica bien sean manuales, conlleven el uso de objetos, maquinaria específica o se centren en la aplicación de técnicas digitales, tienen puntos en común, pues confluyen en acercarse a los sitios arqueológicos desde diferentes ópticas, como pueden ser la geográfica, física, medioambiental, arqueográfica, toponímica o fuentes orales, entre otras.

En los últimos años asistimos a la proliferación de un nuevo tipo de prospección arqueológica de tipo remoto, denominado por nosotros en otro lugar como "prospección informática" (Carrero Pazos *et al.* 2014), por cuanto se trata de estudiar, mediante técnicas diversas de teledetección, sensores remotos, o técnicas de visualización microtopográfica derivadas del tratamiento informático de datos digitales, zonas del territorio que posteriormente se comprueban en campo.

Desde el año 2015, España cuenta con datos LiDAR accesibles públicamente y de forma gratuita a través del Centro Nacional de Información Geográfica (CNIG). El proyecto PNOA-LiDAR[14] proporciona, en formato de libre acceso, datos procedentes de dos coberturas de vuelo LiDAR, una correspondiente al período 2009-2015, y otra desde el 2015 hasta la actualidad (**Figura 5**).

[13] Reflejo de esto es la gran cantidad de trabajos que se publican anualmente sobre estos temas, tanto en monográficos especializados como aquellos orientados al público en general (véase, por ejemplo, Wiseman *et al.* 2007; Kammermans *et al.* 2014; Parcak 2021), y revistas específicas (tal es el caso de *Remote Sensing* o *Archaeological Prospection*).

[14] https://pnoa.ign.es/web/portal/pnoa-lidar/presentacion

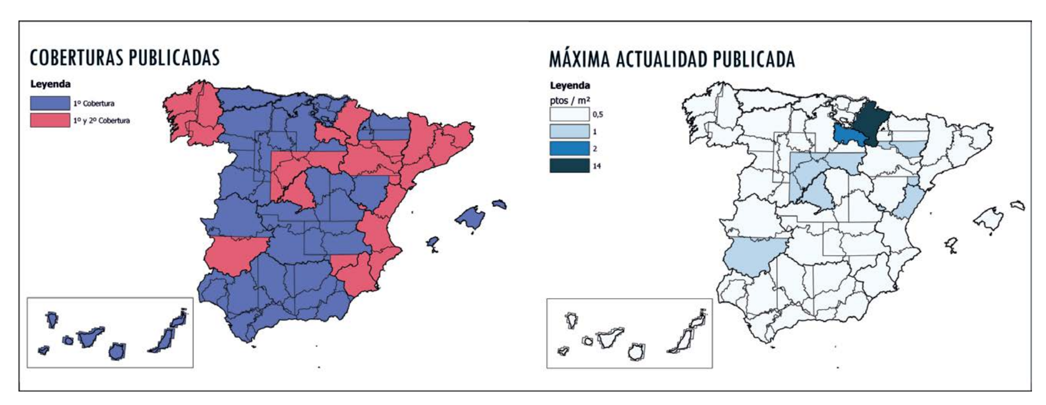


Figura 5. Estado actual del proyecto PNOA-LiDAR, con datos referidos a las dos coberturas disponibles y su publicación, año 2020 (Copyright PNOA, IGN).

El LiDAR (acrónimo inglés de *Light Detection and Ranging*) es un sensor láser activo que genera pulsos y permite realizar mediciones y cálculos gracias a la emisión de un haz de luz sobre la superficie terrestre. Forma parte de todo un conjunto de herramientas de teledetección o sensores remotos que suelen dividirse, en función de su tipología analítica, en aquellos usados sobre elementos terrestres (*Terrestrial Laser Scanning*, TLS), o sistemas aéreos (*Airborne Laser Scanning*, ALS) (Crutchely, Crow 2009).

Emplazado en una avioneta, helicóptero, o dron, etc., este sensor emite pulsos láser que barren el terreno transversalmente, de forma que, al regresar al sensor, recogen información tanto de la superficie terrestre (vegetación, estructuras, etc.), como del tiempo que tardan los pulsos en ser enviados y recibidos, o su intensidad.

En la actualidad, esta tecnología es una de las principales herramientas para la prospección arqueológica, hasta el punto de ser considerada por algunos investigadores como la innovación más importante en adquisición de datos arqueológicos desde la invención del GPS (Bewley *et al.* 2005), por la capacidad para obviar la vegetación de la superficie del suelo (Bennett *et al.* 2011, 2012).

Los estudios y ensayos relativos al desarrollo de los sistemas láser se iniciaron en la década de los sesenta del siglo XX, pero no será hasta finales de los noventa cuando dicha tecnología sea aplicada a la teledetección, así como a proyectos de desarrollo de datos geoespaciales, que hasta entonces se centraban en el uso de información altimétrica (producción de ortofotos y ortoimágenes), la generación de curvas de nivel, sostenidos en Modelos Digitales del Terreno (San José Albacete 2011). El desarrollo de técnicas analíticas que utilizan datos LiDAR ha experimentado un notable impulso debido, en gran medida, a la versatilidad de aplicaciones.

En el caso de la Arqueología, el LiDAR comenzó a ser utilizado en la primera década de siglo XXI, si bien algunos autores ya anticipaban su uso para la detección de fosas, muros y otro tipo de construcciones (por ejemplo, Ackermann 1999). Esta tecnología se ha conformado como una herramienta muy potente y útil en campos como la prospección y Arqueología espacial, especialmente por la generación de Modelos Digitales del Terreno de gran precisión (Bennett *et al.* 2011, 2012; GDAL Development Team 2014).

El análisis del territorio a partir de estos datos proporciona avances importantes, dado que nos permite observar anomalías superficiales que pueden ser comprobadas en el campo, desde pequeños túmulos hasta castros, afloramientos naturales o construcciones de diferente tipología, tanto para zonas de monte bajo como bosques o áreas de vegetación densa (Carrero Pazos *et al.* 2014). Además, el LiDAR permite la revisión de referencias arqueológicas antiguas que necesitan ser actualizadas, pudiendo sobrepasar el análisis informal basado en la fotointerpretación. Su uso se muestra fundamental cuando, por ejemplo, revisamos los catálogos patrimoniales de los organismos oficiales que fueron creados paulatinamente a partir de los años ochenta en el contexto de la construcción de las grandes infraestructuras estatales y urbanísticas de nuestro país (por ejemplo, Rodríguez del Cueto, Carrero Pazos 2021). En términos temporales, aquel momento podría denominarse como una “era pre-informática”, usando, como hemos comentado, herramientas cartográficas manuales a la hora de cartografiar sitios arqueológicos, como la triangulación mediante brújula. Esta circunstancia provocaba la existencia de errores en la localización de los yacimientos que, a veces, podían llegar a superar el kilómetro, tal y como se ha tenido ocasión de comprobar en el caso, por ejemplo, del Megalitismo del Noroeste peninsular (véase al respecto Ferreira de Sousa 2013; Rodríguez del Cueto, Carrero Pazos 2021). En este sentido, la comparación de los datos oficiales con una cartografía de alta resolución obtenida a partir de datos LiDAR permite mejorar la precisión en la localización de los sitios y, en consecuencia, la creación de mapas de distribución espacial del patrimonio arqueológico mucho más exactos.

1.3.1. El tratamiento de los datos LiDAR

Los datos LiDAR en bruto, denominados frecuentemente con el término nube de puntos (en un formato específico, LAS o LAZ), deben ser tratados computacionalmente para que resulten de utilidad en Arqueología. Estos datos pueden clasificarse según su altura, número de retorno o intensidad, entre otras características, y además puede seleccionarse qué vegetación (media, alta o baja) se obvia o mantiene. En Arqueología es frecuente, por ejemplo, elegir sólo los puntos correspondientes al terreno y eliminar la vegetación y estructuras artificiales. De esta forma, el usuario puede generar un Modelo Digital de Elevaciones de gran precisión que, obviando la vegetación y aplicando toda una serie de filtros analíticos, permite detectar estructuras arqueológicas en superficie.

En la **Figura 6** puede observarse una nube de puntos en la que se muestran los puntos del terreno (en marrón) y diferentes tipos de vegetación, desde baja y media hasta zonas de arbolado. Esta distinción entre diferentes elementos vegetativos ha sido posible porque la nube de puntos ha sido sometida a un proceso de clasificación. Las claves para la clasificación de puntos LiDAR fueron definidas, en su momento, por la Sociedad Americana de Fotogrametría y Teledetección (ASPRS).[15] Este trabajo de clasificación es necesario y técnicamente está orientado a asignar, a cada punto, un valor numérico que se corresponda con el tipo de objeto que reflejó el pulso láser (por ejemplo, las carreteras son codificadas con el número 11, mientras que la vegetación baja con el número 3).[16]

[15] https://www.asprs.org

[16] En la actualidad, el CNIG ofrece, para los datos LiDAR de segunda cobertura, una versión ya clasificada, que permite a los investigadores ahorrarse el paso de la clasificación. No obstante, realizar el trabajo de clasificación es muy recomendable, puesto que los resultados analíticos (la creación del MDT y posteriores visualizaciones digitales) dependen, en último término, del proceso de filtrado de los datos y su clasificación.

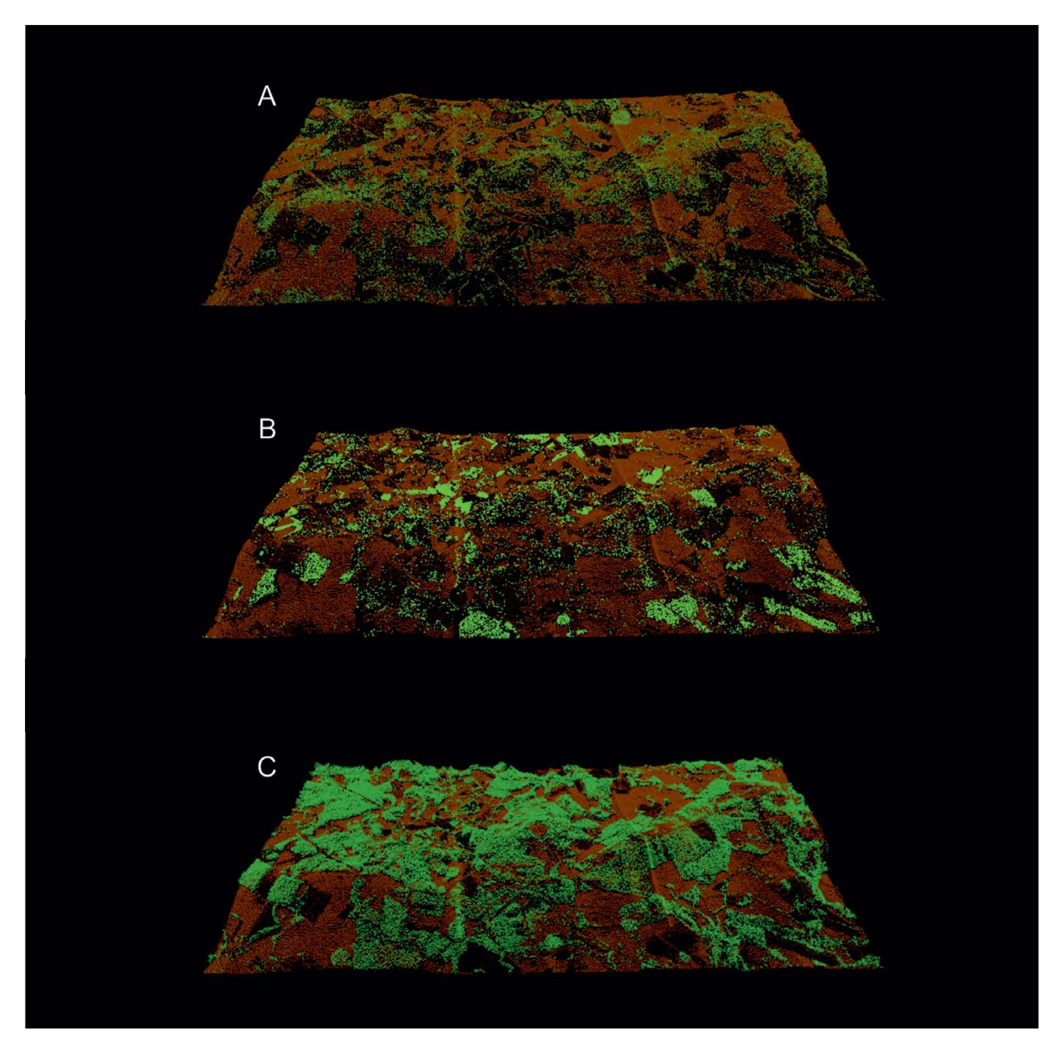


Figura 6. Nubes de puntos LiDAR clasificadas por tipos de vegetación. **A:** Puntos correspondientes al terreno y vegetación baja. **B:** Puntos correspondientes al terreno y vegetación media. **C:** Puntos correspondientes al terreno y vegetación alta.

La densidad de puntos de las nubes proporcionadas por el CNIG depende del vuelo seleccionado y la zona de estudio, aunque por norma general se suele contar con densidades de 0,5 puntos/m^2 para la primera cobertura, y de 0,5-0,4 puntos/m^2 en la segunda, pudiendo llegar a números mayores en áreas determinadas. Los modelos digitales de elevación que podemos obtener a partir de estas densidades de puntos cuentan, pues, con elevada resolución, y hay que tener en cuenta que, en el caso de investigaciones espaciales, estructuras artificiales como caminos, taludes o carreteras serán observables en el MDE y, por lo tanto, influirán en los cálculos que se puedan derivar de aquí, por ejemplo, en un análisis de visibilidades.

A continuación, procedemos a describir el flujo de trabajo a la hora de producir MDEs derivados de datos LiDAR. Aunque hemos señalado que hoy es frecuente utilizar la versión





Figura 7. Caso de estudio. Castro de Viladonga (Castro de Rei, Galicia), perteneciente a la cultura castreña de la Edad del Hierro del Noroeste. **A**: Ortoimagen procedente del PNOA. **B**: Vista aérea del sitio arqueológico (copyright *wikipedia commons*).

de datos LiDAR ya clasificada descargable a través del CNIG, incluimos este paso analítico por considerarlo de relevancia en las investigaciones arqueológicas.

A efectos de caso de estudio, hemos seleccionado uno de los castros más paradigmáticos del interior gallego, como es el de Viladonga (Castro de Rei, Lugo, coordenadas geográficas: 43.1609665606058, -7.3887993922510145). Su localización sobre un pequeño promontorio y sus prominentes estructuras perimetrales nos permitirán proporcionar, a través del manejo de los datos LiDAR, visualizaciones que realzan la microtopografía de la zona y permiten observar mejor la estructura y planta del yacimiento.

1.3.2. Creación de un Modelo Digital del Terreno a partir de datos LiDAR

Para conseguir un Modelo Digital del Terreno sin vegetación (*bare-earth*, en su terminología inglesa), podemos utilizar, entre otros, el software *LASTools. LASTools*[17] es, en realidad, un conjunto de librerías que permiten el tratamiento de datos LiDAR mediante la ejecución de diferentes *scripts* o paquetes analíticos. Es un software que puede ser utilizado en su versión *stand-alone* (*open access*, aunque con límites) o a través de, por ejemplo, un Sistema de Información Geográfica, como *Quantum GIS*, gracias a su integración en esta plataforma mediante un complemento o *plug-in*.

El trabajo metodológico para conseguir un MDT sin vegetación parte de la herramienta *lasground* de *LASTools*, que permite clasificar la nube de puntos LiDAR entre puntos del terreno (clase 2), y aquellos que no pertenecen al terreno (clase 1, correspondiente a estructuras y vegetación). Posteriormente, *las2las filter* nos permite extraer sólo los correspondientes al terreno, obteniendo, de esta forma, una nube de puntos sin vegetación ni estructuras artificiales (**Figura 8)**.

Gracias al uso de *lasground* y *las2las filter* podemos aislar los puntos pertenecientes al terreno, que ahora debemos interpolar para generar una superficie tridimensional. El objetivo es crear una malla regular de triángulos que normalice los datos en zonas sin puntos, calculando nuevos puntos a partir de los vecinos.

Existen varias formas de realizar esto. El propio *LASTools* nos incluye la herramienta *las2dem*, que crea un Modelo Digital de Elevaciones utilizando una técnica de interpolación lineal estándar, basada en la creación de una red de triángulos irregulares (*Triangular Irregular Network*, TIN en su acrónimo más conocido) que conectan los puntos LiDAR entre sí formando

[17] https://rapidlasso.de/

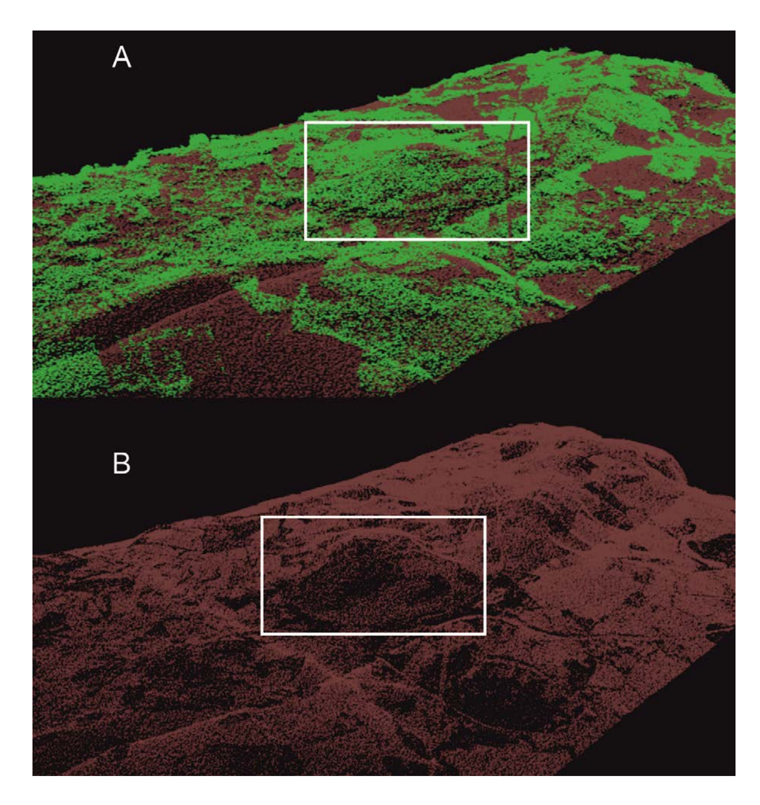


Figura 8. A: Resultados del filtraje de clasificación realizado con el *script lasground* de *LASTools* (puntos del terreno en marrón; estructuras artificiales y vegetación en verde), con factor de exageración 3 para una mejor visualización. **B:** Resultados del filtrado de clasificación realizado con el script *las2las filter* (selección de puntos solamente del terreno). Los cuadros blancos de ambas imágenes identifican el Castro de Viladonga.

la red de triángulos. La versión gratuita de *LASTools* puede suponer, en este caso, un problema, al contar con una capacidad límite de análisis de datos (unos 6 millones de puntos por archivo). Si ese límite es sobrepasado, el software permite analizar los datos, pero en los resultados se crearán unas líneas negras artificiales que impiden una correcta visualización. Este es un procedimiento creado *ad hoc* para que aquellos trabajos que busquen analizar grandes cantidades de datos adquieran la licencia "pro" del software, con un pago justificado en el análisis de un volumen de datos elevado.

No obstante, esto puede solucionarse realizando la interpolación en cualquier Sistema de Información Geográfica al uso. Nosotros recomendamos *SAGA GIS*,[18] uno de los SIG *open access* más potentes para trabajos de base raster. Cargando sencillamente la nube de puntos filtrada (recordemos, aquella que posee sólo los puntos del terreno), podemos realizar diferentes interpolaciones. En términos simples ahora, y recordando lo indicado anteriormente, la interpolación es una técnica matemática que permite estimar valores en zonas no conocidas (por ejemplo, aquellas en las que no existen puntos LiDAR, o su densidad es muy baja), a partir de los valores existentes en las zonas conocidas. Existen diferentes tipos de algoritmos de interpolación, y todos ellos tienen un impacto importante en los resultados; en este caso utilizaremos el *Inverse Distance Weighted* (IDW), que funciona particularmente bien para distribuciones de puntos irregulares. El proceso de interpolación generará nuevos puntos en aquellas áreas donde existían estructuras artificiales y vegetación ya que, recordemos, éstas fueron eliminadas en el proceso de filtraje al quedarnos sólo con los puntos del terreno. Los

[18] https://saga-gis.sourceforge.io/en/index.html

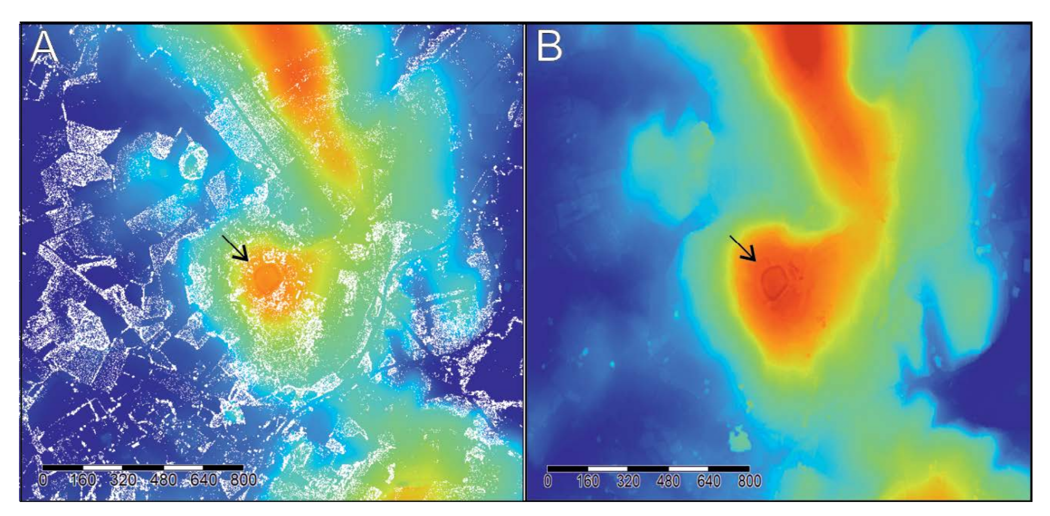


Figura 9. Interpolación de la nube de puntos LiDAR correspondiente al Castro de Viladonga (identificado con una flecha en ambas imágenes), en SAGA GIS. **A**: Nube de puntos sin interpolar. **B**: Interpolación del Modelo Digital del Terreno mediante *Inverse Distance Weighted*.

resultados pueden observarse en la **Figura 9**. Con ello, hemos generado el Modelo Digital del Terreno a partir de datos LiDAR.

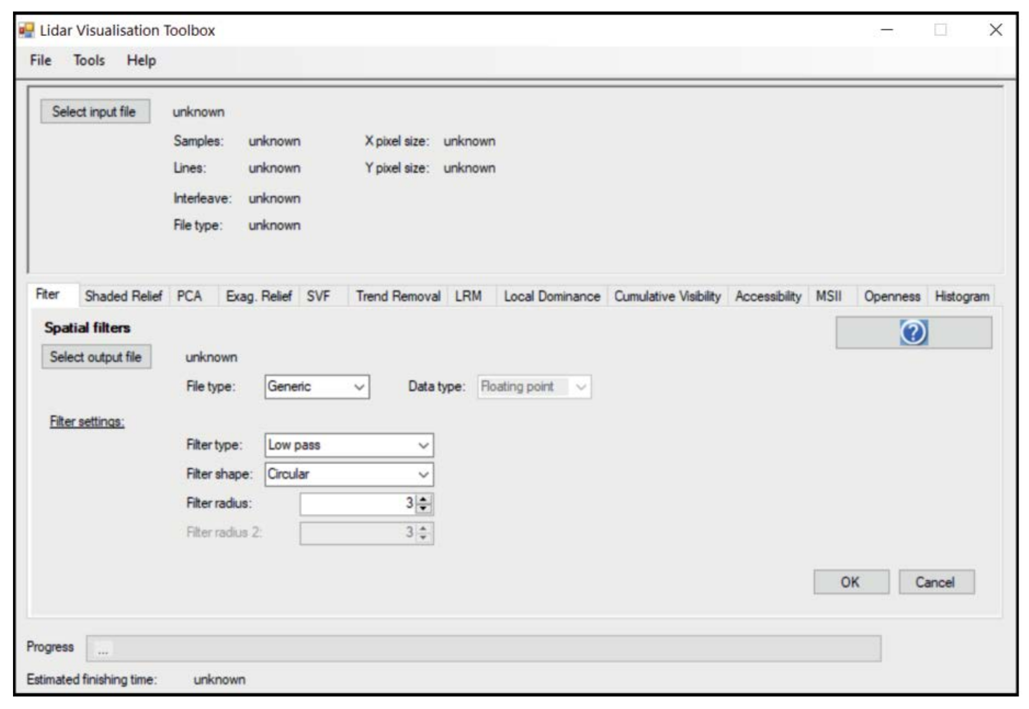
1.3.3. Técnicas de visualización microtopográficas

Los estudios de identificación de elementos y sitios arqueológicos con tecnología LiDAR se encuentran hoy en día en continua proliferación. A la hora de producir imágenes derivadas de datos LiDAR existen muchas opciones, bien mediante sombreados analíticos o a través de la aplicación de algoritmos de cálculo específicos. Tal y como reconocen E. Cerrillo Cuenca y A. López López (2020), la bibliografía europea es muy numerosa en este sentido, y mucha de ella está orientada al estudio arqueológico. Baste con señalar, ahora, el libro de Z. Kokalj y R. Hesse *Airborne laser scanning raster data visualization* (2017), una de las principales aportaciones para visualizaciones derivadas de datos raster obtenidos a partir de LiDAR.[19]

En el contexto peninsular, el Noroeste ha sido una de las zonas en la que, en los últimos años, han aparecido gran cantidad de trabajos que utilizan estos datos, sobre todo con respecto a yacimientos de la Edad del Hierro y época romana (Costa García 2015). En el caso del Megalitismo contamos con trabajos parciales que han ido integrando este tipo de estudios en proyectos más generales (Mañana Borrazás *et al.* 2010; López Romero, Mañana Borrazás 2013), junto a investigaciones integrales que ponen de relieve la potencialidad de estas herramientas (Carrero Pazos *et al.* 2014; Carrero-Pazos, Vilas-Estévez 2016).

Hoy en día, cualquier SIG contiene módulos específicos que permiten generar sombreados analíticos a partir de MDEs, aunque existen algunas herramientas, como el *LiDAR Visualisation Toolbox20* (LiVT) (Hesse 2016) o la *Relief Visualisation Toolbox21* (RVT) (Zakšek *et al.* 2011) que fueron diseñadas para responder a necesidades arqueológicas, como el estudio de estructuras modernas e históricas en el paisaje, o el análisis de cambios en la topografía de terreno (**Figura 10**).

[19] Pueden consultarse también los trabajos de Hesse 2010; Kokalj *et al.* 2011; Štular *et al.* 2012 y Zakšek *et al.* 2011.
[20] https://sourceforge.net/projects/livt/
[21] https://iaps.zrc-sazu.si/en/rvt#v; https://rvt-py.readthedocs.io/en/latest/



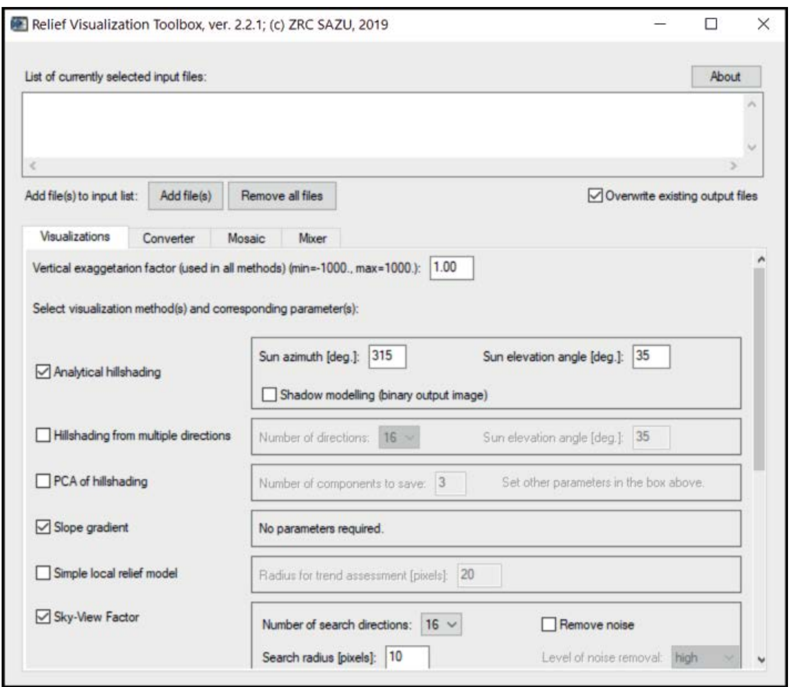


Figura 10. Dos de los softwares más utilizados para la generación de resultados analíticos a partir de datos LiDAR. **Arriba:** *LiDAR Visualisation Toolbox;* **Abajo:** *Relief Visualization Toolbox.*

El libre acceso a estos softwares de procesamiento, tanto en lo que se refiere al tratamiento de datos LiDAR como a los análisis de imagen digital sobre MDEs, ha permitido también su democratización entre usuarios no experimentados, coadyuvando la generación de resultados analíticos que pocas veces conllevan un conocimiento específico previo sobre el funcionamiento de los algoritmos empleados para crear las imágenes resultantes (Cerrillo Cuenca, López López 2020: 223).

a) Analytical Hillshading

El sombreado analítico es, quizás, el método de visualización más común en cartografía digital. Utiliza una técnica de iluminación del modelo digital del terreno con luz direccional, simulando un determinado ángulo de iluminación del sol que puede ser modificado manualmente (Horn 1981; Zakšek *et al.* 2012). En la actualidad la mayor parte de los SIG contienen un módulo para realizar un sombreado analítico,[22] aunque softwares como RVT o LiVT permiten la producción

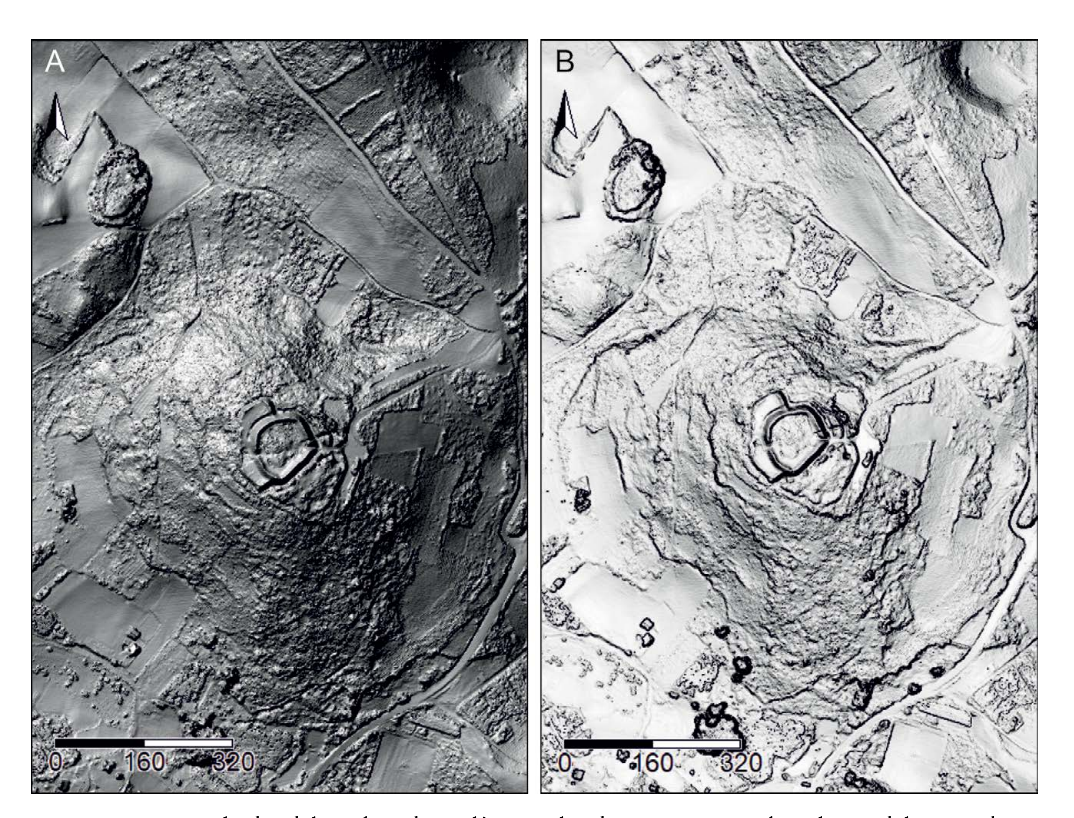


Figura 11. Resultados del sombreado analítico realizado en *SAGA GIS*, sobre el MDE del Castro de Viladonga generado previamente. **A:** *Analytical hillshading*, método standard (Azimut: 315; Altura: 45; Exageración: 4). **B:** *Analytical hillshading*, método combinado (Azimut: 315; Altura: 45; Exageración: 4). En ambos casos, la planta y estructuras murarias del castro son perfectamente visibles.

[22] Siguiendo el ejemplo anterior, *SAGA GIS* contiene el módulo *Analytical Hillshading* (Tarini *et al.* 2006), que permite seleccionar entre varios métodos de sombreado, nombradamente estándar, estándar-máximo 90 grados, sombreado combinado (*combined shading*) (en la línea de Tzelepis *et al.* 2008), trazado de rayos (*ray tracing*) u oclusión ambiental (*ambient occlusion*), que permite producir imágenes en las que las partes del territorio más ocluidas aparecen sombreadas (Landis 2002; Pharr, Green 2004).

de múltiples resultados y, por lo tanto, son muy recomendables para realizar comparativas (**Figura 11**).

b) Uso de filtros espaciales

El filtrado espacial de imágenes es una de las principales técnicas de procesamiento digital encaminado al realce. Se trata de técnicas que comprenden un conjunto de operaciones tendentes a mejorar la calidad visual de una imagen. Pretenden, por tanto, suavizar o reforzar los contrastes espaciales de tal forma que los valores de cada píxel de la imagen se asemejen o diferencien más a los pixeles que los rodean (Aldalur, Santamaría 2002: 31).

Los filtros espaciales se basan en el uso de una ventana analítica para calcular los valores nuevos de una celda central, que frecuentemente se denomina núcleo o *kernel*. Estos filtros son el resultado de la aplicación de una función matemática específica que atañe a los valores de las celdas cubiertas por una ventana (un proceso que, técnicamente, se denomina "convolución"; Burrough, McDonnell 1998). La ventana analítica se desplaza a través de toda la imagen (por ejemplo, un MDT), asegurándose que todas las celdas reciben un nuevo valor filtrado (Conolly, Lake 2009: 260). El filtrado espacial se expresa matemáticamente como

$$C_{ij} = f(\sum_{i-m}^{i+m}\sum_{j-m}^{j+m} Cij \cdot \lambda ij)$$

donde la celda central (*Cij*) es una función (*f*) de cada una de las celdas vecinas (*Cij*), de una ventana de radio *m*, donde el valor de cada celda (*Cij*) se multiplica por una ponderación *λij*, para luego sumar estos valores y determinar el valor de la celda del núcleo. El objetivo es incrementar o decrecer la correlación entre las celdas vecinas, creando filtros de paso bajo (suavizado) o filtros de paso alto (aquellos que resaltan los cambios).

Los filtros de paso bajo suavizan la variabilidad local en un raster, y suelen utilizarse para difuminar la presentación de resultados. Se aplican para eliminar los datos extremos, por ejemplo, en los modelados hidrológicos, o para suavizar las distribuciones de frecuencia de artefactos con propósitos ilustrativos. Los filtros de paso bajo calculan la celda central de la ventana analítica como un promedio de sus vecinos, que puede ser, por ejemplo, la media o la moda. Por otra parte, los filtros de paso bajo también pueden ser utilizados para medir la variabilidad local (Conolly, Lake 2009: 262), aunque aquí se plantea la existencia de un problema de borde, ya que en las celdas cercanas al borde de un mapa raster no se calcula el valor central, porque se calcularía en base a valores erróneos (puesto que son áreas sin datos).

Tabla 1. Ejemplos de filtro de paso bajo con ventanas analíticas de 3x3 y 5x5, que pueden utilizarse en *GRASS GIS (r.mfilter)* creando un archivo de texto con esta estructura.

Filtro de paso bajo		Filtro de paso bajo	
TITLE	3x3	TITLE	5x5
MATRIX	3	MATRIX	5
1 1 1 1 1 1 1 1 1		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
DIVISOR	0	DIVISOR	0
TYPE	P	TYPE	P

Podemos calcular un filtro de paso bajo de forma sencilla con *GRASS GIS*, utilizando el módulo *r.mfilter*, a partir de un filtro creado por el usuario (por ejemplo, un archivo de texto, en formato *txt*), con la estructura reflejada en la **Tabla 1**.

Por su parte, los filtros de paso alto se definen como la antítesis de los filtros de paso bajo. Tienen el efecto de resaltar la variabilidad local, de ahí

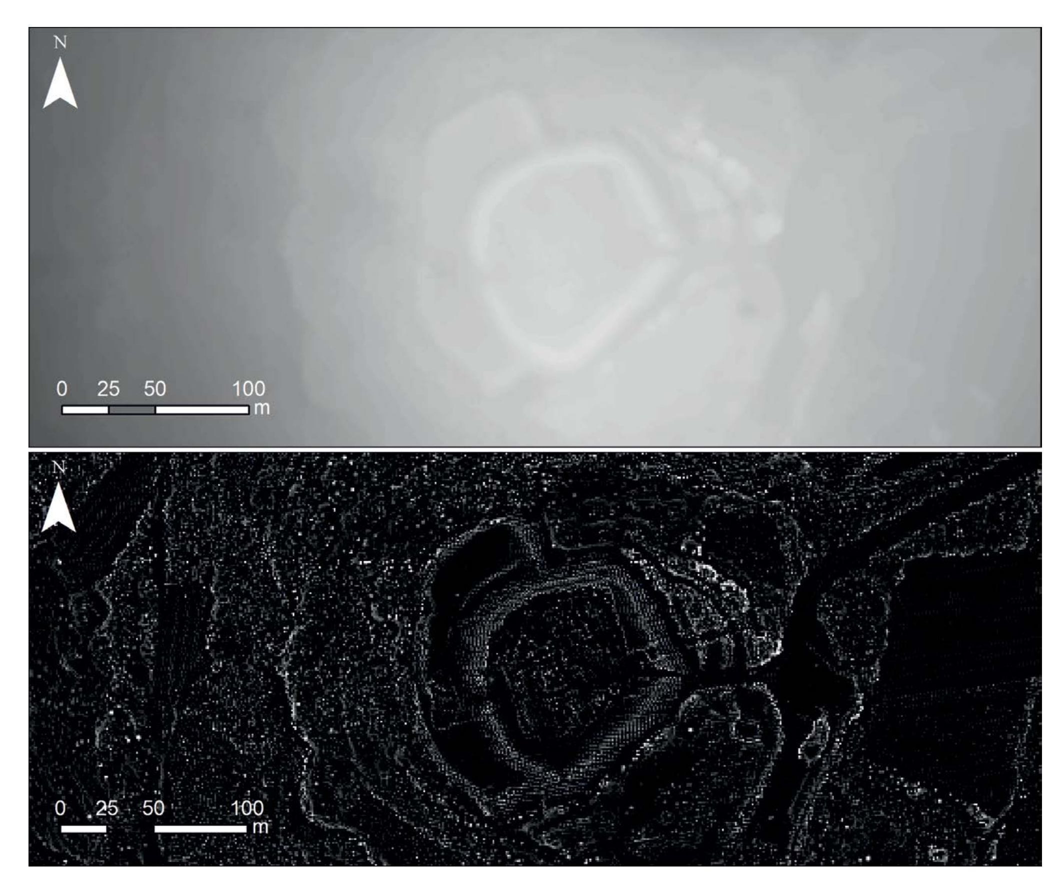


Figura 12. Uso de filtros espaciales para la detección de las estructuras del castro.
Arriba: de paso bajo. **Abajo:** de paso alto.

que en Arqueología estos filtros hayan sido utilizados para detectar bordes topográficos (*edge-detection*); por ejemplo, áreas en las que existen cambios rápidos en la elevación. Así, han sido una de las herramientas más usadas para extraer y vectorizar información de fotografías aéreas antiguas (Forte 1995).

En términos de álgebra de mapas, los filtros de paso alto se definen como el producto de la resta del mapa raster original y el mapa de paso bajo. Recordemos que las técnicas en álgebra de mapas se refieren al conjunto de operaciones que se llevan a cabo sobre un mapa raster (Tomlin 1990), y se suelen utilizar para construir modelos de análisis espaciales (Burrough, McDonnell 1998: 184) (**Figura 12**).

a) Modelos de relieve local (Local Relief Models, LRM)

Los modelos de relieve local son un conjunto de técnicas de visualización que realizan una representación de la superficie topográfica con valores de elevación a escala local (Hesse 2010, 2013, 2016). Este tipo de análisis es muy interesante porque genera un mapa de colores en el que se aprecian formas positivas y negativas en el territorio (áreas cóncavas y convexas),

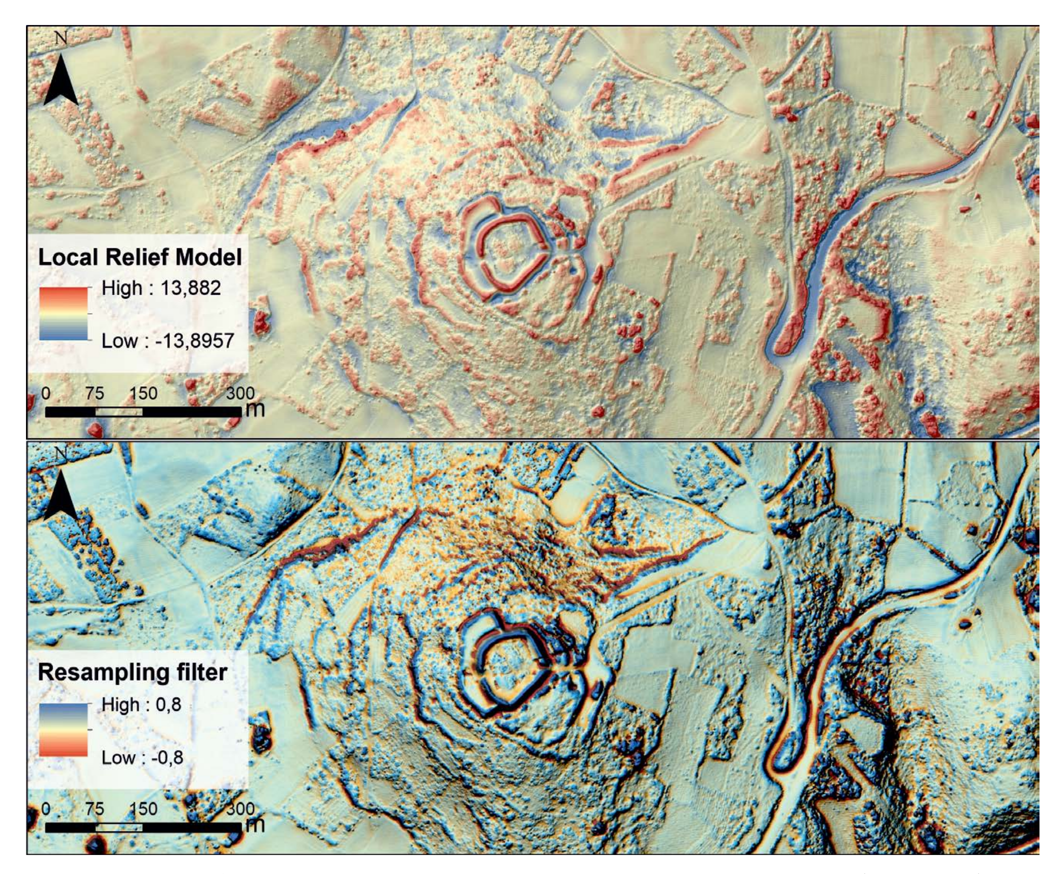


Figura 13. Arriba: Modelo de relieve local realizado con la *toolbox LRM* de *ArcGIS* (Novák 2014). **Abajo:** Filtro resampling de *SAGA GIS*. En ambas imágenes puede observarse la planta del Castro de Viladonga en el centro.

con resultados muy llamativos. La técnica, desarrollada por R. Hesse (2010) utilizando los trabajos de A. Humme *et al.* (2006) y J. K. Hiller y M. Smith (2008), puede aplicarse utilizando el módulo *r.local.relief* de *GRASS GIS*, la *toolbox* LRM que D. Novák (2014) ha puesto a disposición pública para *ArcGIS*, o el *resampling filter* de *SAGA GIS*; módulos, todos ellos, con los que se pueden conseguir productos visuales en los que la microtopografía del terreno, y por ende, las estructuras arqueológicas, aparecen resaltadas (**Figura 13**).

1.3.4. Comparación y discusión de los resultados analíticos

Las técnicas y métodos descritos en las páginas anteriores permiten producir resultados analíticos de indudable valor para un proyecto de prospección e investigación arqueológica, tanto en lo que se refiere a la búsqueda de nuevos sitios arqueológicos, al estudio de estructuras arqueológicas en superficie, como a la transferencia de los resultados al público en general.

A tenor de lo expuesto, no existe una técnica o método más recomendable que los demás, sino que un buen trabajo debe orientarse hacia la discusión de los resultados generados por

todos ellos. Cada sitio arqueológico es susceptible, pues, de un debate específico en torno a qué método es el más adecuado, y esto dependerá siempre de los objetivos iniciales de la investigación. Por lo tanto, es necesario que el usuario, especialista o no, conozca los entresijos analíticos, problemas y consecuencias del uso de un determinado algoritmo o software en pro de otro, porque los resultados pueden diferir sustancialmente.

De lo que no cabe duda es que el tratamiento de datos LiDAR y la generación de imágenes analíticas con objetivos arqueológicos se ha constituido como una de las herramientas más interesantes para la prospección arqueológica en los últimos años, a juzgar por el número de publicaciones que se vienen sucediendo (consúltese al respecto Cerrillo Cuenca, López López 2020). Aunque ciertos autores, no sin razón, refieren que esta tecnología ya no debería ser considerada bajo los calificativos de nueva o novedosa, pues existe desde hace bastante tiempo, lo cierto es que en España sus aplicaciones arqueológicas son recientes. Desde 2011 (y hasta 2019), E. Cerrillo Cuenca y A. López López (2020) dan cuenta de la existencia de 59 trabajos publicados que, con diferente temática y profundidad, utilizan técnicas de visualización analíticas basadas en datos LiDAR. Dos han sido, entre otros, los casos de estudio en los que el LiDAR ha supuesto una verdadera "revolución" en el campo arqueológico. Nos referimos, en primer lugar, a la detección de nuevas evidencias de campamentos romanos (Costa-García 2017; Costa-García, Fonte 2017; Cordero Ruíz *et al.* 2017) y, en segundo lugar, a la detección de túmulos megalíticos en el paisaje (Cerrillo-Cuenca 2017; Cerrillo-Cuenca, Bueno-Ramírez 2019), tanto de forma manual (análisis e interpretación visual de las imágenes analíticas por parte del investigador), como semiautomática (algoritmos que detectan túmulos en el paisaje de forma automática) (Cowley 2012; Berganzo-Besga *et al.* 2021). En el contexto peninsular, el trabajo de I. Berganzo-Besga (*et al.* 2021) es clave, pues a partir de algoritmos de detección automática basados en datos LiDAR, los investigadores han conseguido detectar cerca de 10.000 estructuras tumulares en todo el paisaje gallego, elementos que podrían corresponderse con túmulos prehistóricos. Se trata de una de las primeras aproximaciones que diseña un algoritmo de detección específico y en acceso abierto para detectar estructuras tumulares y, además, por primera vez, se analiza un territorio tan extenso como es la totalidad de la región gallega.

Todo ello nos indica que el campo de la detección automática de sitios arqueológicos sólo ha empezado a despegar y auguramos un futuro muy fructífero, coincidiendo con la paulatina revisión de los catálogos patrimoniales autonómicos que, en no mucho tiempo, permitirán promover estudios espaciales con elevada precisión (por ejemplo, Rodríguez del Cueto, Carrero Pazos 2021).

Concluimos, pues, indicando la necesidad de revisar los métodos actuales de prospección arqueológica relativos a la detección de yacimientos arqueológicos en España. Redefinir las metodologías de prospección arqueológica supone que, indefectiblemente, debamos acudir a la interdisciplinariedad y tratemos de incluir en los equipos arqueológicos a especialistas en tecnologías de teledetección, sensores remotos, uso de drones, o promovamos la formación en estas técnicas y métodos entre los estudiantes de Arqueología.

En este capítulo se han mostrado diferentes posibilidades de análisis digitales que se pueden conseguir al generar un modelo digital del terreno de gran precisión a partir de datos LiDAR, observando de una forma más fidedigna estructuras arqueológicas positivas y negativas en el

territorio. El uso de los modelos tridimensionales y los procesos analíticos comentados han aumentado considerablemente en las últimas décadas, no sólo con objetivos museísticos o mediáticos, sino también como una de las líneas más activas en la investigación arqueológica. Ello quizás sea debido a que, entre otros factores, los métodos y técnicas descritas se nutren de los avances informático-tecnológicos que, con fugacidad, permiten generar nueva información arqueológica al tiempo que salvaguardan aquella de su destrucción.

Capítulo 2.

Arqueología del territorio y paisaje a través de los Sistemas de Información Geográfica

2.1. Las investigaciones con Sistemas de Información Geográfica y estadística espacial

Los Sistemas de Información Geográfica forman parte de las Ciencias de Información Geográfica o Tecnologías de Información Geográfica. Han sido definidos como "la herramienta tecnológica más poderosa aplicable a la Arqueología desde el invento de la datación por radiocarbono" (Wescott, Brandon 1999: 1). Y, sin embargo, desde sus primeras aplicaciones en nuestro campo, las críticas no han sido pocas. Llegadas en mayor medida desde las corrientes historiográficas postprocesualistas, consideraban a los SIG una tecnología "demasiado dependiente de supuestos sobre la importancia de patrones espaciales, en un espacio artificial deshumanizado" (Pickles 1999: 50-52).

Dejando de lado los debates teóricos, en los que nos centraremos más adelante, la realidad es que los SIG han supuesto, desde los primeros momentos, un importante giro metodológico en la disciplina arqueológica, bien por su capacidad analítica de trabajo, bien porque han posibilitado el desarrollo de líneas de investigación que nos permiten estudiar y contrastar, con argumentos cuantitativos, modelos teóricos sobre las sociedades del pasado.

La historia de los Sistemas de Información Geográfica puede retrotraerse a los años 1950, en concreto en países como Gran Bretaña o Estados Unidos. Allí se suceden prontos intentos de desarrollar cartografías asistidas por ordenador, como el *Atlas of the British Flora*,[23] aunque será en los años 1960 cuando las herramientas cartográficas adquieran un importante impulso, primero en Estados Unidos, Gran Bretaña y Canadá. En este sentido, será en el *Laboratory for Computer Graphics and Spatial Analysis* de la Universidad de Harvard donde Fisher empezaría a diseñar y desarrollar un *software* específico para aplicaciones cartográficas, que más tarde se denominaría *Symap* (*Synagraphic Mapping Technique*), el primer programa de creación de mapas en ser comercializado (Comas, Ruiz 1993; Moldes Teo 1995).[24]

[23] O los avances que por aquel entonces se llevaron a cabo en las universidades de Washington e Iowa con Garrison y McCarty al frente, respectivamente (Comas, Ruiz 1993).

[24] Otros dos de los sistemas más relevantes en la evolución de los primeros SIG fueron *Grid* e *Imgrid*. Creados por el *Department of Landscape and Architecture* de la Universidad de Harvard, incluían aspectos de modelado dinámico, y

El caso de Canadá es, quizás, el más interesante de todos, porque es allí donde, en 1966 se desarrollará por primera vez un SIG en el sentido moderno del término (*Canadian Geographic Information System*, gracias al trabajo de Tomlinson).[25]

La iniciativa comercial en Estados Unidos estuvo, y lo está actualmente, monopolizada por el *Environmental Systems Research Institute* (ESRI), que desde sus inicios elaboró sistemas destinados al tratamiento automático de la información geográfica. A destacar es el sistema vectorial *Polygon Information Overlay System*, que sirvió de base para la creación de *Arc/Info* en 1982, una herramienta que integraba el procesado de datos cartográficos e información asociada. Su evolución, *ArcView*, veía la luz en 1992 como un SIG de visualización y consulta aunque sin capacidades de análisis, aspectos incluidos en las versiones posteriores (*ArcGIS*, a partir de 2001).

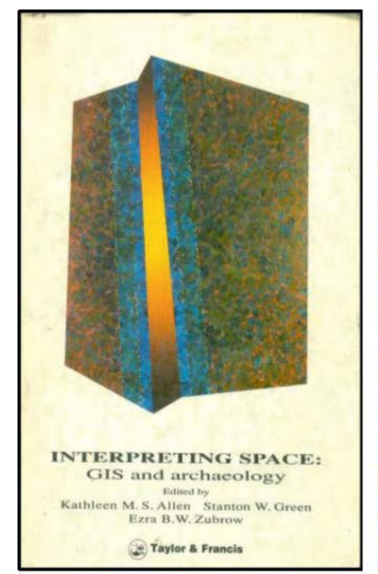
En España, los SIG se desarrollaron en sus inicios gracias a la acción de iniciativas de instituciones y administraciones del Estado. A principios de los años setenta, el Instituto Geográfico y Catastral, actual Instituto Geográfico Nacional, adquirió el *software Symap* norteamericano para desarrollar un Atlas Nacional de España. En años posteriores se diseñaría el *Sistema de Información Geográfica Nacional*, que funcionaría hasta 1980. En los noventa, la gran expansión tecnológica que los SIG experimentan también es sentida en España, con la creación de la *Asociación Española de Sistemas de Información Geográfica* (AESIG).

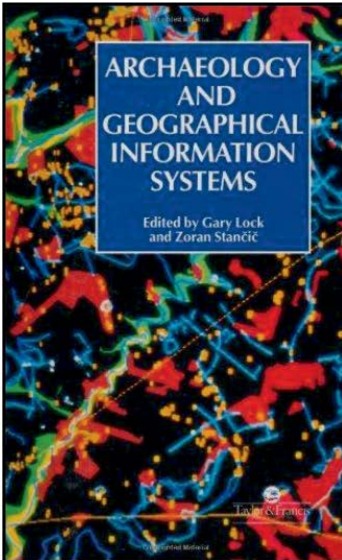
En la actualidad los SIG son utilizados en diferentes campos, como la Geografía, la Arqueología, la Ingeniería o Geología, con presencia en las universidades españolas. En este sentido, los primeros trabajos que utilizaron los SIG en Arqueología se desarrollaron, de nuevo, en Estados Unidos, en la gestión de los recursos culturales y el desarrollo de modelos predictivos de localización de asentamientos (Grau Mira 2006). Sin embargo, fue en la década de los noventa cuando asistimos a la proliferación de diversos monográficos, obras clásicas y de obligada consulta que ponen de relieve la importancia de esta herramienta en la disciplina arqueológica. Nos referimos, por ejemplo, al *Interpreting Space: GIS and Archaeology* (1990) de K. Allen, S. Green y E. Zubrow, así como su análoga europea editada por G. Lock y Z. Stančič, *Archaeology and Geographical Information Systems: An European Perspective* (1995). También podemos destacar el monográfico *Anthropology, Space, and Geographic Information Systems* (Aldenderfer, Maschner 1996) o, ya a fines de siglo, el *Practical Applications of GIS for Archaeologists. A Predictive Modeling Kit* (Wescott, Brandon 1999) (**Figura 14**).

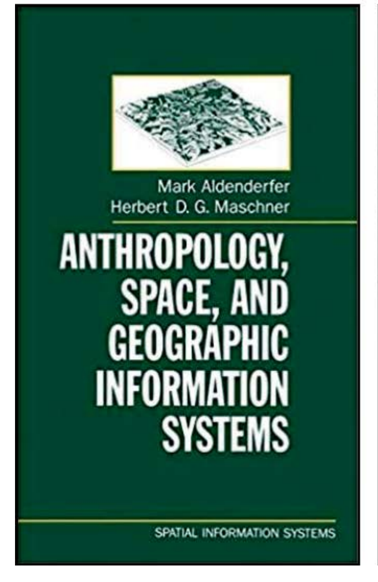
Con el cambio de siglo aparecen toda una serie de manuales específicos orientados a la explicación de las principales técnicas y métodos en trabajos con información geográfica vectorial y raster, y su aplicación sobre casos arqueológicos. Nos referimos, por ejemplo, al *Spatial Technology and Archaeology. The archaeological applications of GIS*, de D. Wheatley y M.

resulta interesante señalar que *Imgrid* fue uno de los primeros programas SIG que se utilizó en España (1977), por el *Centro de Información y Documentación de Santander* (Moldes Teo 1995). Desde entonces, las administraciones públicas desarrollaron diferentes programas de cartografía asistida por ordenador con objetivos de almacenaje y representación de datos geográficos. Así es como el *United States Geological Survey* (USGS) desarrolló en 1973 el *Geographical Information and Retrieval Analysis System* (GIRAS) para gestionar y analizar información sobre los recursos del suelo de manera eficiente.

[25] En el caso de Gran Bretaña, más concretamente en el ámbito universitario, se desarrollaron nuevas líneas de investigación para el análisis espacial. Los primeros proyectos corrieron a cargo del *Institute of Terrestrial Ecology*, de Huntingdon, en el cual se utilizó el *Power Samas Card Calculating System*, un *software* con la finalidad de crear mapas digitales de las distribuciones vegetales del territorio nacional (Dabezies Damboriarena 2009).







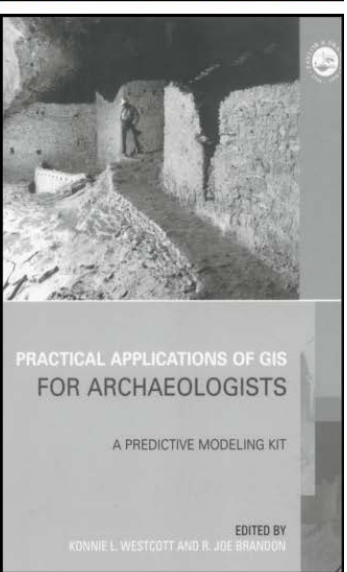


Figura 14. Monográficos sobre SIG aplicados a Arqueología publicados la década de 1990.

Gillings (2002) o al que es, hasta la fecha, el monográfico más completo sobre Sistemas de Información Geográfica en Arqueología: *Geographic Information Systems in Archaeology*, de J. Conolly y M. Lake (2006; versión en castellano, 2009). Quince años después de su publicación, este manual académico sigue siendo la principal síntesis sobre el tema en Arqueología.

Otros trabajos que podemos citar para estos momentos son el *GIS and Landscape Archaeology,* de H. Chapman (2006), o los orientados al desarrollo de casos de estudio específicos, como el *GIS and Archaeological Site Location Modelling*, de M. Mehrer y K. Wescott (2006) (**Figura 15**).

A partir de 2010 asistimos a un proceso de renovación en la Arqueología computacional, con la proliferación de estudios y proyectos que rescatan viejos procedimientos analíticos y los

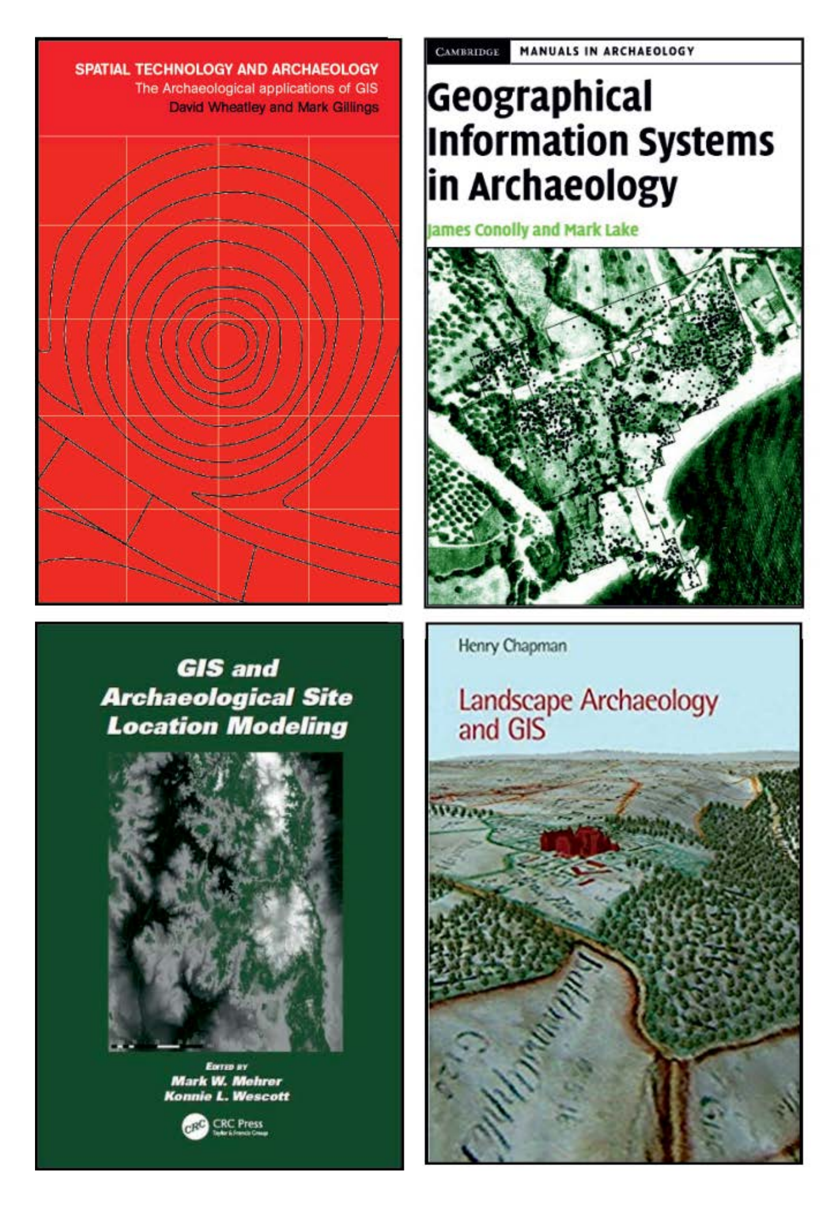


Figura 15. Monográficos sobre SIG aplicados a Arqueología, publicados a inicios del S. XXI.

re-orientan a través de técnicas nuevas, caso de la simulación estadística, como parte del giro cuantitativo que la Arqueología parece estar recuperando en la actualidad (Kristiansen 2015).

Se suceden trabajos que incorporan herramientas estadísticas avanzadas junto a renovadas reflexiones sobre el uso del SIG y sus métodos en casos arqueológicos, como el *Computational Approaches to Archaeological Spaces*, editado por A. Bevan y M. Lake (2013), el *Computational Approaches to the Study of Movement in Archaeology*, de S. Polla y P. Verhagen (2014), o el *Archaeological Spatial Analysis. A methodological guide*, editado por M. Gillings, P. Hacigüeller y G. Lock (2020). Destaquemos, también, el monográfico *Modelling Human Behaviour in Landscapes*, de O. Nakoinz y D. Knitter (2016), en su totalidad desarrollado mediante el lenguaje de



Figura 16. Últimos monográficos de referencia en el campo de la Arqueología computacional, desde 2010.

programación *R Statistics*, que anticipa, a la vez que ejemplifica, el uso de ciencia abierta y transparente en Arqueología (**Figura 16**).

En lo que se refiere a la Arqueología española, la incidencia y desarrollo fue menor en los primeros años en comparación con otros países, a juzgar por el reducido número de síntesis. El monográfico "Los SIG y el análisis espacial en Arqueología", coordinado por J. Baena Preysler (*et al.* 1997) abría un camino que sería continuado por investigadores puntuales, como M. Á. Moreno Gallo, con su tesis doctoral "Megalitismo y Geografía. Análisis de los factores de localización espacial de dólmenes de la provincia de Burgos" (2004). Destaquemos, asimismo, el uso de estas herramientas en la gestión de información arqueológica, como ponen de relieve los trabajos de S. Fernández Cacho (2009) en el marco de la arqueología patrimonial andaluza, o los trabajos del *Incipit* en el marco de la Arqueología del paisaje gallega (Parcero Oubiña, González Pérez 2007).

El monográfico de L. García Sanjuán, *Introducción al Reconocimiento y Análisis Arqueológico del Territorio* (2005), marcó un punto de inflexión en la aplicación de las tecnologías geoespaciales en la Arqueología española y, en concreto, en lo concerniente a las técnicas de prospección. Este trabajo supuso el contrapunto español al contexto europeo del momento, dejando patente que muchos de los proyectos de prospección de universidades y centros de investigación nacionales se situaban, también, a la vanguardia de la ciencia arqueológica.

Todo esto coincidiría, años después, con la publicación en castellano del monográfico de J. Conolly y M. Lake, *Sistemas de Información Geográfica aplicados a la arqueología* (2009). Un manual que, paulatinamente, fue siendo introducido en los planes académicos. La Arqueología computacional comenzaba, de esta forma, a sentar base para un prolijo futuro académico.

2.2. Un camino pedregoso: las críticas del postprocesualismo y el determinismo ambiental

La evolución historiográfica de los trabajos con SIG y estadística espacial en Arqueología coincide, también, con la aparición de voces críticas desde los inicios, llegadas principalmente desde corrientes historiográficas postprocesualistas y tradicionales (Wheatley 1993; Thomas 1993, 2004), que veían a los SIG como una suerte de "venganza positivista" (Taylor 1990: 211-212).

Una de las críticas esgrimidas que mayor impacto historiográfico y académico ha obtenido es la referida al determinismo ambiental, entendiendo como tal la dependencia de estas tecnologías en supuestos sobre la importancia de patrones espaciales en espacios artificiales, en los que las sociedades humanas parecen no estar presentes (Pickles 1999: 51-52). Las argumentaciones son bastante más complejas (véase Gaffney, Van Leusen 1995; Hacigüzeller 2012 para un análisis más extenso) pero, en términos generales, el determinismo ambiental de los trabajos basados en los SIG es palpable, argumentan los críticos, en la importancia conferida a factores como la topografía, tipos de suelo o vegetación en la conformación y estructuración de las sociedades pasadas, dejando de lado factores menos accesibles como pueden ser determinados rasgos sociales o rituales. C. Tilley (2004: 218) apunta hacia este último sentido cuando, al investigar los paisajes megalíticos, señala que éstos no pueden ser entendidos sólo desde el punto de vista de los mapas, los diagramas, fotografías o descripciones, que son, en último término, representaciones o modelos de la realidad. De hecho, como representaciones que son, argumenta, no logran, ni lograrán transmitir nunca, una comprensión corporal de los restos y sitios prehistóricos. Para él, los SIG y las simulaciones estadísticas no pueden substituir a las experiencias humanas de "sentir" los lugares arqueológicos.

Esta línea de críticas en torno al determinismo ambiental y los problemas de "representacionismo" ha coincidido, también, con contra-argumentos por parte de investigadores funcionalistas. Muchos han sido quienes, por ejemplo, consideran que las representaciones base de un trabajo con SIG, un mapa o argumento estadístico, son aproximaciones realistas porque permiten modelizar los restos físicos y la dimensión "humanista" del espacio (véase al respecto, por ejemplo, Lake, Woodman 2003). Lo cierto es que, como indica P. Hacigüzeller (2012: 256), los investigadores que utilicen estas herramientas deben ser conscientes de que están trabajando con modelos. En consecuencia, deben reflexionar acerca del proceso de creación del conocimiento que están realizando, y optar por modelos de pensamiento que alberguen un discurso crítico en torno a los modelos

creados. Porque, continúa la autora (Hacigüzeller 2012: 257), caminar a través de un paisaje o excavar un determinado yacimiento arqueológico no es la única forma legítima de crear conocimiento arqueológico, y de la interacción con la tecnología o de la reflexión sobre los modelos analíticos que se puedan construir sobre el pasado también se crean narrativas históricas que son perfectamente válidas.

2.3. Cuestiones técnicas de inicio sobre los SIG

En el *I Encuentro Internacional de Informática Aplicada a la Investigación y la Gestión Arqueológica* (Martín de la Cruz *et al.* 2004: 28), M. Toledano Rubio define a los SIG como "un sistema de hardware, software y procedimientos diseñados para realizar la captura, almacenamiento, manipulación, análisis, modelización y presentación de datos referenciados espacialmente para la resolución de problemas complejos de planificación y gestión". Esta definición puede completarse con las funcionalidades específicas que J. Conolly y M. Lake (2009: 30) describen con relación a las investigaciones arqueológicas (**Figura 17**):

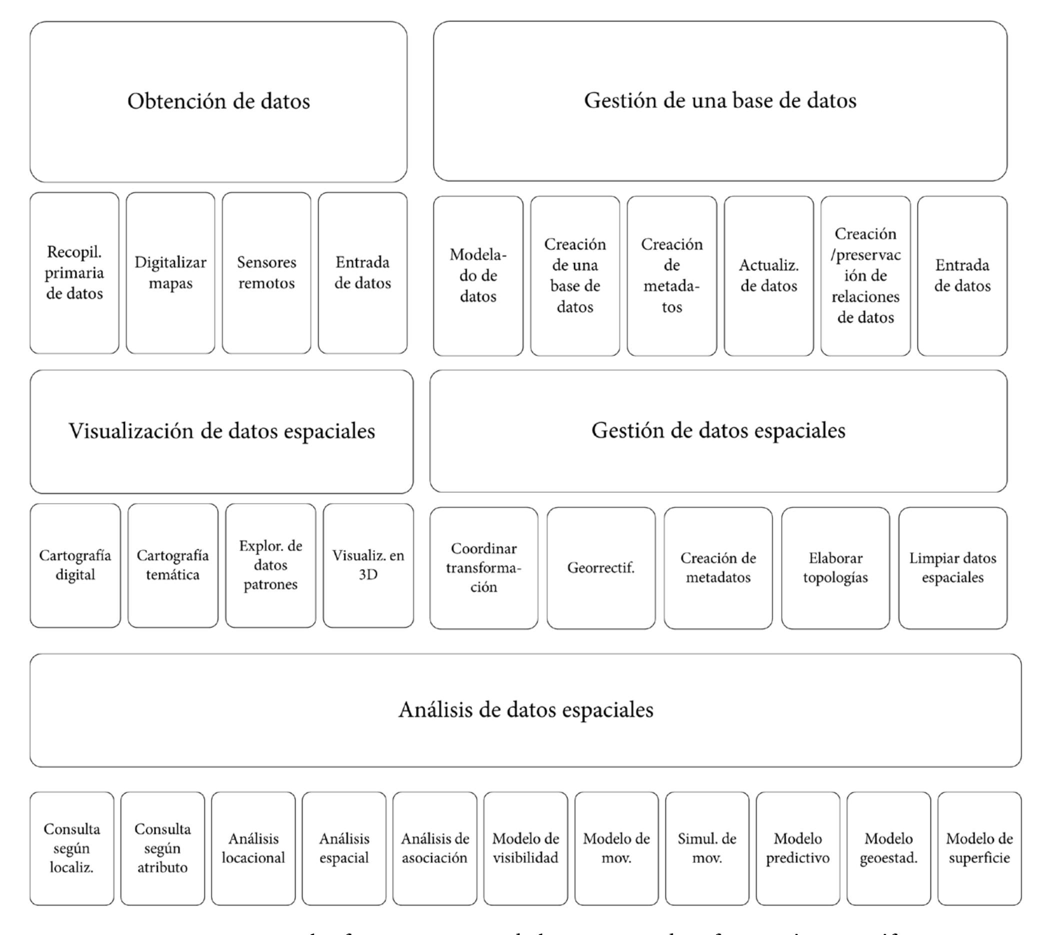


Figura 17. Principales funciones y usos de los Sistemas de Información Geográfica (modificado de Conolly, Lake 2009).

Una de las aplicaciones más frecuentes de los SIG en las investigaciones arqueológicas redunda en la creación de mapas de distribución arqueológica, tanto desde perspectivas *intrasite* (escala de yacimiento) como *offsite* (escala de territorio). Los SIG permiten trabajar con datos estadísticos, su cartografía y transformación en variables activas para estudios más profundos. Nos permiten, en definitiva, explorar patrones espaciales partiendo del análisis de las distribuciones arqueológicas, con el objetivo de llegar a los procesos o decisiones que los han conformado.

A grandes rasgos, los SIG funcionan con dos tipos de datos: raster y vectorial. Los datos vectoriales son aproximaciones orientadas a objeto, por lo que la realidad es representada como puntos, líneas y polígonos. Por otro lado, las superficies raster son campos de datos continuos en los que la información es descrita mediante series de píxeles con determinados valores[26] (**Figura 18**).

La representación de objetos espaciales mediante datos vectoriales y raster nos permite generar modelos que pueden componerse de objetos (a menudo, polígonos o líneas), redes (por lo general puntos, interrelacionados entre ellos), y campos (habitualmente imágenes). Todos ellos se convierten en capas temáticas al ser incluidos en un Sistema de Información Geográfica, aunque debe existir un trabajo previo de planificación y definición, observando qué información reflejan, para qué van a ser utilizadas, y cuál es la mejor forma de representarlas.[27]

El uso de modelos vectoriales cuenta con ventajas importantes. Entre las primeras está la precisión espacial y su resolución, que es independiente. Además, este tipo de datos presentan una alta facilidad de asociación con tablas de atributos, lo que nos permite gestionar a la

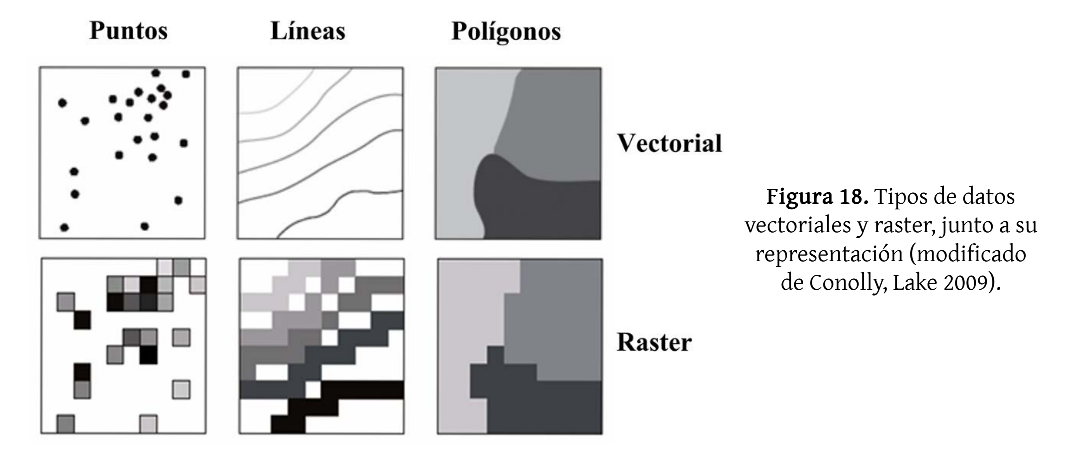


Figura 18. Tipos de datos vectoriales y raster, junto a su representación (modificado de Conolly, Lake 2009).

[26] Los modelos raster se componen de una red de celdas que, en términos generales, son cuadradas y mantienen un tamaño uniforme. Una fotografía o un modelo digital de elevaciones son ejemplos comunes de datos raster. Cada celda puede expresar números continuos, enteros, rangos o categorías.

[27] También existen versiones 3D de los modelos de datos vectoriales y raster que suelen utilizarse en Arqueología. Así tenemos, por ejemplo, para datos vectoriales, la representación mediante una Red de Triángulos Irregulares (TIN). Se trata de un modelo sólido, compuesto por una red alámbrica de triángulos interconectados entre sí que generan una superficie sólida. Por otra parte, los modelos 3D raster pueden obtenerse, por ejemplo, mediante imágenes de rayos X, y se componen de píxeles volumétricos (frecuentemente denominados *voxels*).

vez información documental y espacial de forma fluida. Sin embargo, como desventajas señalaremos que suelen requerir una importante demanda computacional.

Trabajar con un modelo raster supone hacerlo con un tipo de datos muy expresivos a la hora de representar fenómenos variables pero continuos, como la elevación, o aquellos con límites difusos, caso, por ejemplo, de los cambios en la vegetación. Sin embargo, de entre sus desventajas principales está la resolución, que siempre es fija, y por lo tanto limitada a la hora de asociar información documental mediante tablas de atributos.[28]

Un aspecto importante es la cuestión de la escala. Cuando estudiamos objetos espaciales o pretendemos explicar una realidad espacial concreta del mundo que nos rodea, bien sea mediante información vectorial o raster, la escala supone uno de los elementos más relevantes a considerar. En Arqueología debemos tener en cuenta que la escala de nuestros análisis debe adaptarse al fenómeno que vayamos a estudiar porque de ello dependerán, por un lado, los datos que utilizaremos y, por otro, el engranaje analítico a desarrollar.

Las investigaciones arqueológicas a menudo utilizan series de mapas vectoriales y raster como base formal para sus productos cartográficos, es decir, como soporte. No obstante, mayor importancia debemos conferir a la cartografía analítica, entendida como una fuente de información y no solamente como base cartográfica. En concreto, nos referimos a los modelos de superficie, un conjunto de herramientas analíticas que busca representar y modelizar las características de la superficie terrestre actual. Destacamos, así, el Modelo Digital de Elevaciones (MDE) (denominado también Modelo Digital del Terreno (MDT)), y el Modelo Digital de Superficie (MDS). Estamos ante representaciones cartográficas del territorio actual, y son simplificaciones que no están exentas de errores.

Una de las técnicas matemáticas más utilizadas en cartografía es la interpolación. Se trata de la derivación de una superficie continua a partir de un conjunto de observaciones individuales (Conolly, Lake 2009: 131), y la selección de la técnica de interpolación más adecuada va a depender de la estructura de los datos muestrales, y también del resultado que se busca y las características del modelo de superficie. Es una técnica que realiza una predicción de datos basada en observaciones circundantes (Burrough, McDonnell 1998), pero sólo puede ser aplicada a fenómenos espaciales cuantitativos que posean un mínimo grado de estructura, es decir, cuando los valores de las observaciones presenten algún tipo de correlación. Este es el caso, por ejemplo, de los datos de elevación; recordemos la interpolación realizada sobre una nube de puntos LiDAR en capítulos anteriores. Cuanto más cerca estén dos lugares en el paisaje, más similar será su elevación, por lo que la interpolación conseguirá crear una superficie más exacta.[29]

[28] Hasta no hace mucho tiempo era necesario realizar una elección previa al comienzo de un proyecto con SIG, pues los softwares tendían a ser mayoritariamente vectoriales o raster. Hoy en día la situación ha cambiado y la mayor parte del software disponible permite trabajar y analizar de forma profunda y rápida tanto datos vectoriales como raster.

[29] En aquellos casos en los que existe una correlación débil, como por ejemplo una distribución de artefactos en el paisaje, es frecuente recurrir a métodos de interpolación de la geoestadística, que valoran la variación espacial en un conjunto determinado de datos, como el *kriging*.

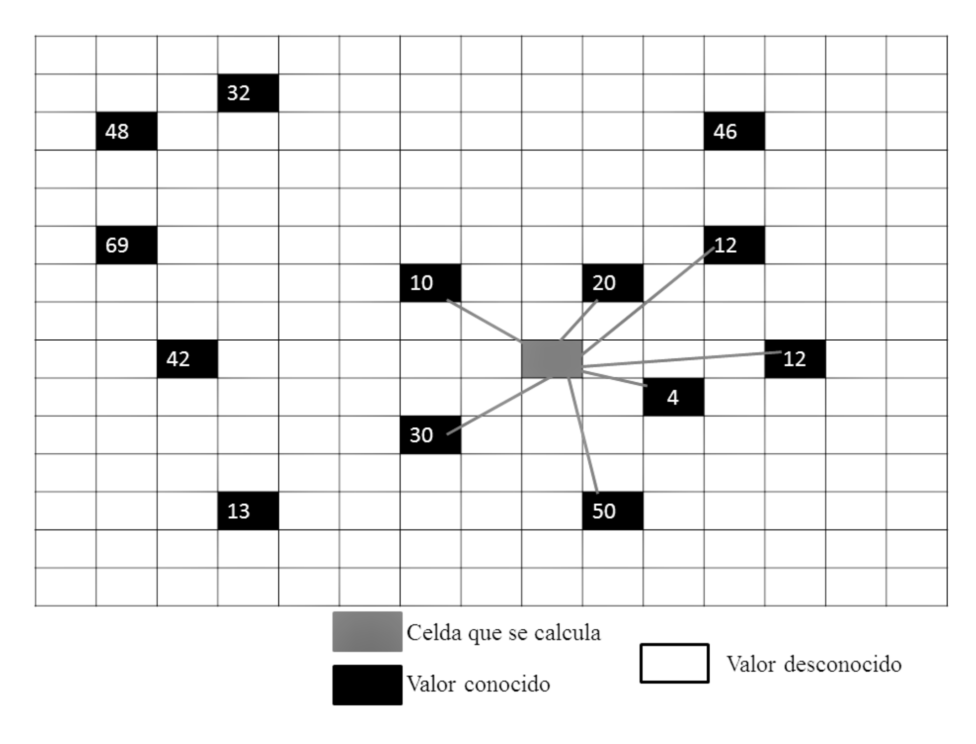


Figura 19. Ejemplo esquemático del proceso de interpolación con distancia media ponderada (IDW) de 4 vecinos más próximos.

Los SIG nos proporcionan, en la actualidad, herramientas analíticas con las que realizar una interpolación de manera sencilla.[30] La interpolación lineal equivale, de forma conceptual, a trazar una línea recta entre dos puntos con valores conocidos. Cualquier punto de esta línea puede pronosticarse a partir del establecimiento de su distancia relativa entre ambos puntos. Sin embargo, tomar sólo dos puntos en una recta puede no suponer una predicción exacta, por lo que a menudo se utiliza la distancia media ponderada de un número determinado de vecinos más próximos, denominada ponderación de la distancia inversa (*Inverse Distance Weighted*, IDW), que puede establecerse entre un radio de entre 4, 8 u 12 vecinos más próximos, lo que influirá en la creación de la superficie final (**Figura 19**).

Los *Splines*, por su parte, son métodos indicados para superficies que varían lentamente, como las elevaciones (Conolly, Lake 2009: 140). Pueden definirse como representaciones matemáticas, en forma de rejilla, que unen determinados puntos para trazar superficies curvadas. Lo que nos interesa es que estos métodos de interpolación se adecúan bien a los datos originales, y conservan la variabilidad de los datos cuando nos movemos en escalas pequeñas.

[30] Estas cuestiones esconden, sin embargo, bastante más complejidad, pues existen modelos globales que permiten establecer tendencias generales en los datos, pero fallan al predecir valores de variación locales. En Arqueología, el método de interpolación más utilizado es el análisis de superficie de tendencia (Conolly, Lake 2009: 133), que acopla una superficie matemática a una distribución espacial de valores de atributos cuantitativos, es decir, representa un proceso espacial (véase por ejemplo Hodder, Orton 1976: 155-174; Bove 1981). Por otro lado, es necesario indicar los modelos locales, que permiten obtener una información más exacta de la variabilidad local. Destacaremos sólo dos de los más utilizados, como son el *Inverse Distance Weighted* (IDW) y los *Splines.*

Sin lugar a duda, el modelo de superficie más utilizado es el Modelo Digital de Elevaciones (MDE), un mapa raster en el que cada celda o píxel que lo compone tiene un valor que representa la elevación del terreno en ese punto. Un MDE puede ser creado a partir de fuentes primarias o secundarias. Las primarias se refieren a datos extraídos, por ejemplo, de forma manual durante el trabajo de campo, utilizando GPS diferenciales o mediante LiDAR, tal y como hemos realizado en capítulos anteriores. Las fuentes secundarias son, por ejemplo, curvas de nivel (hipsografías vectoriales) o puntos altimétricos digitalizados.

Los MDE son fundamentales en cualquier trabajo que incluya análisis espacial con SIG a escala de paisaje. La elección de su resolución y, por ende, precisión, depende siempre de los objetivos y del área de estudio que se considere. Además, debemos tener en cuenta que se trata del *modelo de un modelo*, porque los contornos sobre los que creamos el MDE son sólo una muestra o interpretación simple del terreno real. No obstante, es cierto que en la actualidad contamos con la ventaja de poder acceder a una fuente de datos cartográficos muy precisa, como son los productos derivados de la tecnología LiDAR; datos digitales a través de los que podemos crear modelos digitales de superficie de elevada resolución y detalle.

Los modelos digitales de elevación son, también, una fuente básica sobre la que conseguir modelos ecológicos y ambientales y, por lo tanto, son el eje vertebrador de los trabajos en Arqueología computacional del territorio y paisaje.

Capítulo 3.

Modelos computacionales, distribuciones, patrones y procesos de puntos[31]

3.1. Modelos computacionales en Arqueología

El uso de modelos de simulación computacional en Arqueología no es reciente. Con una dilatada tradición, los primeros se desarrollaron de forma pionera en la década de los 1970, sufriendo un importante renacimiento en los 1990s y expansión a partir del cambio de siglo (Lake 2014).

En Arqueología, un modelo o simulación computacional puede ser entendido como la simplificación de una situación particular, objeto, paisaje, proceso o dinámica social que se utiliza para mejorar nuestro conocimiento sobre la realidad que representa. Es decir, se trata de la representación simplificada de un sistema que puede usarse para comprender o predecir su comportamiento (Epstein 2008; O'Sullivan, Perry 2013).

Los modelos científicos pueden ser de muy diverso tipo, conceptuales, físicos, matemáticos, empíricos o de simulación. Frecuentemente en forma de modelos computacionales (por ejemplo, de base agente), las simulaciones arqueológicas han sido utilizadas para entender el comportamiento humano a largo plazo, analizar dinámicas observadas en el registro arqueológico, como procesos posdeposicionales o incerteza cronológica (Rubio-Campillo *et al.* 2012; Crema 2015), o evaluar hipótesis a través de simulaciones de las condiciones del pasado (Shennan *et al.* 2013; Balbo *et al.* 2014).

Si quisiésemos estudiar, por ejemplo, la movilidad de ciertos grupos neolíticos peninsulares en un determinado territorio, una buena forma de empezar es analizando las dinámicas estructurales del paisaje en el que ese grupo humano estuvo asentado. La movilidad, además, puede ser entendida en clave de grupo o como un comportamiento individual. Suele tener implicaciones culturales y sociales como la identidad o pertenencia a un grupo, con lo que

[31] Ciertas partes de este capítulo han sido publicadas recientemente en el trabajo Carrero-Pazos, M. (2023). "Análisis de patrones espaciales de puntos para el estudio de tendencias locacionales en distribuciones de yacimientos arqueológicos". *Vegueta, Anuario de la Facultad de Geografía e Historia* 23(1): 97-114.

habrá que diseñar estrategias de análisis para entender patrones espaciales y dinámicas de transmisión de la información (Anthony 1990; Kelly 1992; Burmeister 2000).

De entre las aplicaciones más importantes del uso de modelos de simulación en Arqueología destacamos la evaluación de hipótesis. Tal y como reconoce X. Rubio-Campillo (2017: 54), una simulación permite modelar directamente procesos de interacción social y, de manera aún más relevante, hacerlo como dinámicas espaciales y temporales. Esta habilidad se complementa por el hecho de que, al ser un modelo formal, es posible comparar los resultados de la simulación con la evidencia arqueológica de manera directa a través de métodos cuantitativos.

3.2. Cuantificación y estadística espacial

La cuantificación y la estadística espacial son dos de las ramas de las matemáticas más extendidas en los estudios arqueológicos, y son la base de la cual partir para generar modelos de simulación en Arqueología.

La estadística espacial suele utilizarse para entender mejor un conjunto de datos y sus atributos arqueológicos. El concepto de análisis espacial engloba diferentes áreas, comenzando por la manipulación de datos espaciales (Tomlin 1990), el análisis descriptivo y exploratorio (Bailey, Gatrell 1995; Fotheringham *et al.* 1999) o las simulaciones estadísticas, que buscan determinar, por ejemplo, cuándo los datos observados pueden ser representados por un modelo estadístico conocido (Ripley 1977, 1981; Diggle 1983; Cressie 1991). Por último, la modelización espacial engloba el desarrollo de modelos para predecir el comportamiento espacial, muy utilizados en geografía humana por ejemplo para predecir las tendencias en la preferencia por ciertos lugares u optimizar la localización de las empresas (Wilson 2000).

3.2.1. Una historiografía difusa

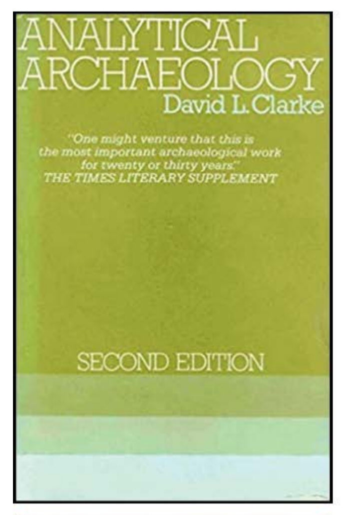
Uno de los momentos más importantes para la disciplina arqueológica se inició a partir de los años sesenta del siglo XX, con el desarrollo de la Arqueología procesual o Nueva Arqueología, que se presentaba con el objetivo de superar a la Arqueología tradicional y otorgar a la disciplina un estatuto científico. No se trató de un fenómeno aislado, sino que esta preocupación funcionalista estuvo presente también en otras ciencias sociales, como la Geografía, la Historia o la Antropología.

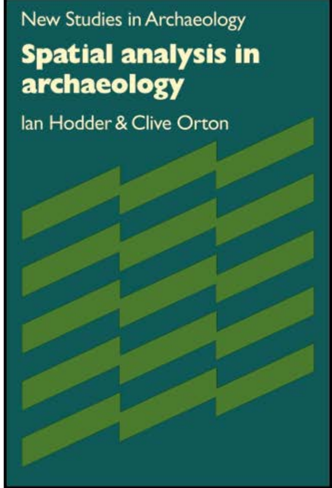
Desde posicionamientos positivistas, la crítica de la subjetividad inherente a las disciplinas humanísticas coincidió con la necesidad de generar conocimientos validados por la práctica empírica, a partir del uso del método hipotético-deductivo. En Arqueología, la praxis debía partir de una idea previa que dirigía la investigación, es decir, un modelo analítico (Orton 1988). El énfasis en el rigor metodológico garantizaba la superioridad de dicho conocimiento científico.

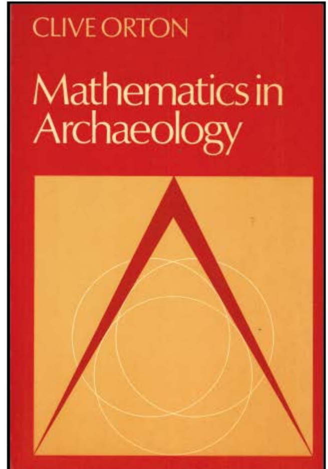
La búsqueda de la objetividad y cientificidad llevó a la Nueva Arqueología a dotarse, entre otras cosas, del desarrollo de modelos matemáticos y estadísticos con los que poner en marcha nuevas interpretaciones arqueológicas, que además venían propugnadas por el rápido avance de la informática. Esto se manifestó, por ejemplo, en el ámbito de la seriación arqueológica, donde la introducción de técnicas matemáticas permitía alcanzar mayor rigor científico (Schiffer 1976: 266).

Uno de los pioneros que resaltó la importancia de la estadística y las matemáticas en arqueología fue D. L. Clarke, que, en su libro *Analytical Archaeology* (1977), sentaba las bases de una nueva corriente metodológica en Arqueología, como era la Arqueología analítica, que incluía técnicas de las disciplinas geográfica y cibernética. También a fines de la década de los setenta, P. Ihm publicaría su *Statistik in der Archäeologie* (1978), donde ya su sugerente título nos presenta una obra que versa sobre estadística en Arqueología. J. E. Doran y F. R. Hodson publican también, por estas mismas fechas, su clásico *Matematics and Computing in Archaeology* (1975), y un año más tarde llegaría el *Spatial Analysis in Archaeology*, de I. Hodder y C. Orton (1976).

Ya en la década de los ochenta tenemos las aportaciones de C. Orton, con *Mathematics in archaeology* (1988, original de 1980), escrito con el objetivo de demostrar que las matemáticas podían explicar muchos de los conceptos e ideas manejados en la Arqueología. S. Shennan, en su *Quantifying archaeology* (1988) introducía los análisis estadísticos complejos, haciendo uso de la cuantificación de los datos arqueológicos (**Figura 20**).







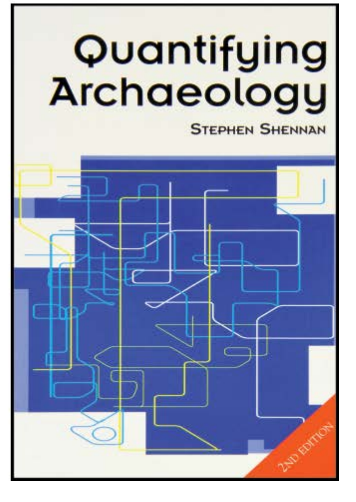


Figura 20. Monográficos de finales del S. XX sobre matemáticas y estadística aplicadas a Arqueología.

Sin embargo, la explosión llegaría en los años noventa con la celebración de reuniones como el *Computer Applications and Quantitative Methods in Archaeology* (Lockyear, Rahtz 1991), que continúan en la actualidad, o para el caso español, la publicación del monográfico "Métodos analíticos y su aplicación a la arqueología", de M. Bandera y F. Chaves Tristán (1994).

Será en estos momentos, a finales del S. XX, cuando asistamos al verdadero despegue de la aplicación de análisis estadísticos y espaciales en Arqueología. Aparecen síntesis como *Digging Numbers. Elementary statistics for Archaeologists* (Fletcher, Lock 1991), *Exploratory Multivariate Analysis in Archaeology* (Baxter 1994), *Statistics in Archaelogy* (Baxter 2003) o *Statistics for Archaeologists* (Drennan 2009). Estos trabajos compendian las principales técnicas y métodos estadísticos con ejemplos aplicados a casos arqueológicos.

En el ámbito español, el *I Encuentro Internacional. Informática Aplicada a la Investigación y la Gestión Arqueológicas*, organizado por J. C. Martín de la Cruz y A. Mª. Lucena Martín en 2003, llamaría la atención sobre la importancia de la informática en Arqueología, aunque será la acción individual de diferentes investigadores la que consiga desarrollar grupos de trabajo enfocados en la aplicación de metodologías cuantitativas en Arqueología. Es el caso, entre muchos otros, de J. A. Barceló, director del Laboratorio de Arqueología Cuantitativa y Aplicaciones Informáticas (UAB), con una extensa bibliografía sobre técnicas matemáticas y métodos computacionales en arqueología (Barceló 1991, 1997, 2007, 2008; Barceló *et al.* 1998; Barceló, Bogdanovic 2014; Barceló, Del Castillo Bernal 2015).

3.2.2. Análisis de datos exploratorio

Cuando en 1976 I. Hodder y C. Orton escribieron su libro "Análisis espacial en Arqueología", dejaban claro que la investigación que les precedía carecía de visiones sintéticas y de una metodología que ofrecer a los investigadores que quisiesen incluir estudios espaciales en sus trabajos. En su opinión, de alguna forma se necesitaba eliminar la subjetividad implícita en la interpretación arqueológica, para lo que veían necesario "desarrollar métodos más rigurosos de interpretación" (Hodder, Orton 1976: 14), basados en la planificación y desarrollo de modelos estadísticos que contemplasen todas las posibilidades de análisis, o la aleatoriedad como hipótesis a contrastar.

Partiendo de la premisa de que un conjunto de datos espaciales puede ser estudiado con relación a toda una serie de atributos (en nuestro caso arqueológicos), existen diferentes tipos de datos, como los nominales y ordinales, que pueden ser estudiados con diferentes técnicas. Los primeros son simples categorías, por ejemplo, tipos de artefactos o yacimientos. Por su parte, los datos ordinales son valores que mantienen un significado de orden, por ejemplo, las fases temporales de un yacimiento. Podemos tener, asimismo, la dicotomía intervalo/ratio. El intervalo tiene un origen arbitrario o respecto a un punto de referencia, mientras que el ratio se define como la relación cuantificada entre dos magnitudes, que refleja su proporción.

A la hora de analizar las variables de una muestra de estudio podemos optar por análisis univariantes, que buscan comprender el comportamiento de una sola variable; bivariantes, que se utilizan para estudiar las relaciones del comportamiento entre dos variables, o los multivariantes, que buscan encontrar patrones observando las relaciones del comportamiento entre múltiples variables a la vez.

Por otra parte, es frecuente el uso de descriptores gráficos que nos ayuden a entender mejor una determinada muestra de estudio. Existen varias formas de mostrar los resultados de los análisis citados anteriormente, como los diagramas de barras, los diagramas de caja y bigotes, que comparan los valores de un intervalo o ratio o los diagramas de dispersión, que comparan los valores medidos de dos variables sobre el mismo objeto. También podríamos destacar los gráficos ternarios, que comparan los valores medidos de tres variables del mismo objeto o los histogramas, que se utilizan para mostrar números que representan rangos de diferente tamaño.

Estos gráficos nos pueden sugerir distribuciones uniformes (todos los eventos son igualmente probables), normales o gaussianas (los eventos son el resultado de un efecto acumulativo de otros eventos, que son independientes uno del otro) y *Poisson* o aleatorias (los eventos son aleatorios, no existe autocorrelación o tendencia en los datos).

3.2.3. Inferencia estadística

Los métodos estadísticos o pruebas de significación sirven para explorar las relaciones en una distribución de datos. Se suelen dividir en test paramétricos y no paramétricos. Los primeros asumen una distribución normal de la muestra, mientras que los segundos se utilizan cuando la distribución podría no presentar normalidad. Para verificar la condición de normalidad de la muestra de datos podemos optar por observar un histograma o realizar una prueba estadística, como el *W de Shapiro-Wilks.*

De entre los test paramétricos, citaremos aquí el *T-Student*, por ser uno de los más utilizados, y de entre los no paramétricos, el *chi-cuadrado* de Pearson, el *test de Wilcoxon-Mann-Whitney* o su extensión *Kruskall-Wallis*, y *test de Kolmogorov-Smirnov*.

El test *T-Student* sirve para establecer la probabilidad de que dos muestras procedan de dos poblaciones distintas. Por su parte, el test de chi-cuadrado pretende encontrar la relación, de existir, entre la frecuencia de una variable y frecuencias teóricas previamente establecidas, o la relación entre dos variables cualitativas establecidas a partir de una tabla de contingencia. Para llevar a cabo el chi-cuadrado necesitamos crear una tabla de contingencia donde se registre el número total de observaciones en contraste con las categorías de los datos. El estudio de los datos se lleva a cabo contrastando los valores observados con los valores teóricos o frecuencia esperada, es decir, valores que cabría esperar si los datos no siguieran una pauta determinada a priori, si no existieran diferencias entre las variables que intervienen en la clasificación (Shennan 1988).

Otra prueba que suele utilizarse para comprobar las diferencias entre dos muestras de datos ordinales es el de *Wilcoxon* o *Mann-Whitney* y su extensión *Kruskall-Wallis* para tres o más grupos. Ambas son pruebas no paramétricas mientras que la segunda es idéntica al *Anova*, con los datos reemplazados por categorías.

El test de *Kolmogorov Smirnov* (test de K-S, en adelante) se utiliza de la misma forma que los anteriores, aunque para realizarlo es necesario convertir los datos en una distribución de frecuencia acumulada (*empirical cumulative distribution function,* ECDF), porque el test mide la diferencia máxima entre las distribuciones de las dos categorías y la compara con la prevista si las muestras se extrajesen de una misma distribución.

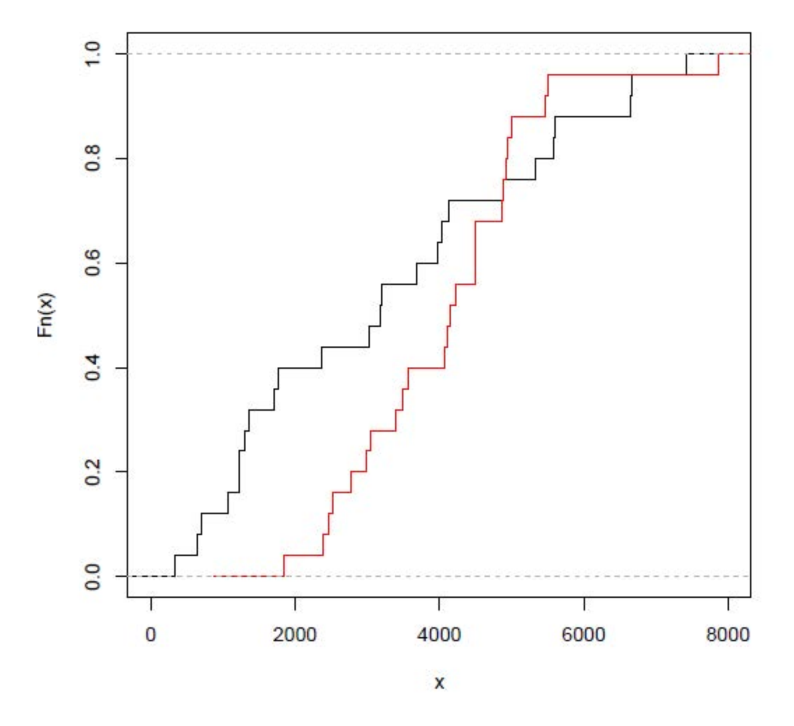


Figura 21. Ejemplo de distribución de frecuencia acumulada de la visibilidad de dos muestras de yacimientos arqueológicos (a partir de Lake, Woodman 2003).

En la **Figura 21**, se observa la diferencia entre una distribución observada (línea roja) y una distribución teórica (línea negra), que nos proporciona el estadístico de K-S (α), y deberá ser comparado con una tabla de valores críticos para saber si es significativo o no, algo que los softwares comunes nos proporcionan directamente.

3.2.4. Hacia una mayor significación: la Aleatoriedad Espacial Completa y la Simulación de Monte Carlo

La estadística espacial tiene como objetivo fundamental el análisis de una distribución de puntos de un determinado proceso, estudiando la ubicación de una serie de eventos de una región concreta para determinar si la muestra se encuentra distribuida de forma aleatoria o no (Bivand *et al.* 2013). Quizás surja la duda de que al analizar una distribución de yacimientos arqueológicos estos nunca estarán dispuestos de forma aleatoria ya que ante todo son creaciones humanas, y por lo tanto su localización responderá siempre a un conjunto de decisiones determinadas. Pese a ser esta una de las asunciones más importantes que debamos admitir, sin embargo, no siempre tiene por qué ser así, y podemos encontrarnos con situaciones en las que existe un patrón de distribución general pero no local, y observar que ciertos yacimientos se localizan siguiendo determinadas pautas generales pero que no cuentan con un patrón local definido, o es aleatorio. Para profundizar en la constatación de si una distribución de yacimientos es fruto del azar o no, se debe estudiar la distribución de partida y compararla con la Aleatoriedad Espacial Completa (*Complete Spatial Randomness*, CSR). En este sentido, una distribución aleatoria se caracteriza por el hecho de que los puntos que la conforman presentan una intensidad homogénea y uniforme a lo largo de toda un área de estudio.[32]

[32] En la mayor parte de los casos, con sólo observar un mapa de distribución ya podemos descartar la aleatoriedad espacial de la muestra de estudio. Algunos de los métodos estadísticos que hemos descrito en el apartado anterior, como la prueba de K-S (Baddeley 2010) se utilizan para rechazar formalmente la CSR.

A la hora de comparar la muestra de estudio con una aleatoria debemos prestar atención al uso de la simulación de Monte Carlo. Cuando las poblaciones de muestra son grandes y diversas, no parece claro cómo generar una muestra arbitraria con la que comparar la distribución real, ni cuántas muestras aleatorias deben tomarse o hasta qué punto una determinada muestra aleatoria es representativa de la población objeto de estudio. La simulación de Monte Carlo reduce esa incertidumbre tomando muchas muestras arbitrarias, por lo general 999 (Conolly, Lake 2009: 400).

El método se consolidó con el desarrollo de la informática a partir de la segunda mitad del siglo XX, y se concretó en la construcción de un rango de distribuciones aleatorias, (comúnmente conocido por su término en inglés, *envelope*) que representan las circunstancias de las que se piensa que la distribución real, observada, se desviará. La mayor parte de los autores (Manly 1991; Fischer *et al.* 1997; Drennan, Peterson 2004; Bevan *et al.* 2013; Bevan 2020), utilizan un número de 999 simulaciones como conjunto de observaciones aleatorias admisible sobre el que aceptar o rechazar la hipótesis nula, a partir de su comparativa con la muestra de datos observada. Así, por ejemplo, si tenemos una muestra poblacional de 121 sitios arqueológicos, la simulación de Monte Carlo permite generar 999 simulaciones aleatorias de 121 puntos cada una, creando un total de 120.879 predicciones que serán comparadas con la muestra arqueológica. De esta forma, comparando ambas distribuciones (la arqueológica y el conjunto de las aleatorias), estaremos en condiciones de observar si el estadístico de los yacimientos cae dentro, en el borde o fuera del rango de valores de las muestras aleatorias, y, por lo tanto, rechazar o aceptar la hipótesis nula de la aleatoriedad en los datos (**Figura 22**).

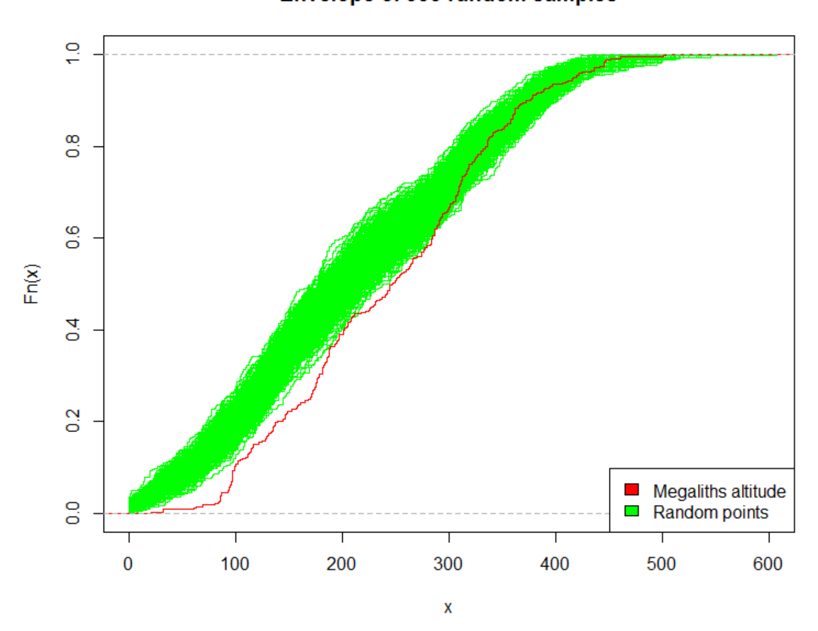


Figura 22. Ejemplo de aplicación de la simulación de Monte Carlo para estudiar la altitud de los monumentos megalíticos de Costa da Morte, Galicia (muestra de 334 monumentos). En verde puede observarse el rango aleatorio (*envelope*), compuesto de 999 simulaciones aleatorias con 334 puntos cada una. En rojo, la distribución observada (monumentos megalíticos).

3.3. Análisis espacial de patrones de puntos

La distribución de puntos en un mapa es siempre una simplificación de un conjunto de eventos complejos a lo largo del tiempo (Bevan *et al.* 2013), y su disposición en el espacio conforma lo que se conoce como patrón de puntos. El análisis de patrones de puntos (*Point Pattern Analysis*, también conocido por sus siglas, PPA) es un campo con larga tradición en Arqueología (Dacey 1973; Whallon 1974), aunque recientemente ha adquirido un renovado impulso gracias a la disponibilidad de técnicas basadas en la simulación estadística con flujos de trabajo intensos (Eve, Crema 2013: 271).

Dentro de esta línea de investigación, los modelos de procesos de puntos (o *Point Process Models*, PPM) se refieren al conjunto de técnicas estadísticas que pueden utilizarse para analizar la distribución y el comportamiento de un determinado conjunto de puntos. La caracterización de un patrón de puntos es de interés para su descripción, y se hace a través de análisis estadísticos y descriptores que intentan definir su estructura.

Para analizar un patrón de puntos se debe asumir, en primer lugar, que la estructura espacial del patrón es el resultado de un proceso puntual no aleatorio. Así pues, los puntos generados son los eventos de este proceso. Se suele medir, por ejemplo, la intensidad de los puntos; esto es, su densidad media (número de puntos esperados por unidad de área), que puede ser constante dentro de una región determinada (proceso homogéneo), o puede variar de localización a localización (proceso no homogéneo), influyendo, de esta forma, en la distribución de los eventos espaciales. Describiendo el tipo de patrón observado se obtiene, a su vez, información sobre el proceso que lo ha generado.

Existen tres tipos de patrones de puntos generados a través de procesos puntuales (**Figura 23**):

- Regular o disperso (*regular*): la densidad es constante y los puntos se disponen alejados unos de otros.
- Aleatorio (*random*): no hay estructura alguna, las posiciones de cada punto son independientes.
- Agrupado (*clustered*): se caracteriza por tener densidades de puntos muy elevadas en ciertas zonas.

Los análisis de patrones de puntos se fundamentan en la comparación entre las propiedades de una distribución teórica aleatoria (bajo condiciones de CSR), y las de la distribución observada. Estas propiedades se concretan en la intensidad del patrón de puntos y en la distinción entre

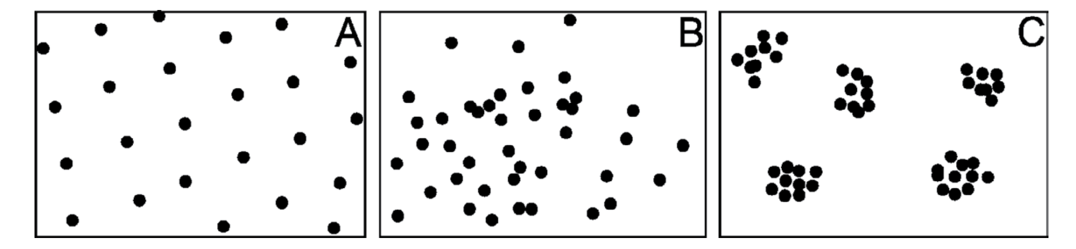


Figura 23. Patrones de puntos. **A:** Regular. **B:** Aleatorio. **C:** Agrupado.

las características que lo componen, denominadas de primer y segundo orden (Bailey, Gatrell 1995: 32-35).

Por norma general, existen, al menos, dos líneas de trabajo en los estudios de las distribuciones de puntos. En primer lugar, aquella que se ocupa del estudio de la forma de la distribución en el espacio y suele estar orientada a estudiar los factores de primer orden. En segundo lugar, aquella que intenta dilucidar cuál es la interacción entre los puntos, conocidos como factores de segundo orden. Ambas líneas forman parte de los análisis de patrones de puntos, un campo que en la disciplina arqueológica está conociendo una importante expansión en los últimos años a juzgar por el elevado número de publicaciones sobre el tema (por ejemplo, Orton 2006; Bevan *et al.* 2013; Palmisano 2014; Visentin, Carrer 2017; Knitter, Nakoinz 2018; Negré *et al.* 2018; Spencer, Bevan 2018; Carrero-Pazos *et al.* 2019; Bevan 2020; Brandolini, Carrer 2020).

Las propiedades de primer orden asumen que la intensidad de la distribución es constante o uniforme a lo largo de la zona de estudio (estacionaria, en términos estadísticos) (Olaya 2014: 294). De esta forma, podemos medir la importancia de las características de primer orden, aquellas que describen la intensidad de los puntos en toda la región estudiada y suelen corresponderse con variables medioambientales, y los efectos de primer orden, que se concretan en aquellos procesos o fenómenos externos que provocan que la intensidad de la distribución de puntos varíe entre diferentes zonas del área de estudio (Bevan *et al.* 2013: 31). En otras palabras, los efectos de primer orden son aquellas observaciones que, en una misma área de estudio, varían de un lugar a otro debido a cambios provocados por factores externos, normalmente medioambientales (O'Sullivan, Unwin 2003: 36).

Para la modelización y control de los efectos de primer orden se suele partir del trabajo con Sistemas de Información Geográfica, creando variables raster a partir de un modelo interpretativo previo que seleccione aquellas que, presumiblemente, puedan haber sido importantes a la hora de explicar el fenómeno histórico en cuestión. En segundo lugar, la regresión logística multivariante es el método que suele utilizarse para seleccionar el conjunto de variables que predice, de forma más precisa, la distribución de datos observada (los sitios arqueológicos) y, con ello, permite controlar los efectos de primer orden (la influencia de las variables medioambientales sobre la distribución de sitios arqueológicos). En capítulos posteriores nos detendremos en las explicaciones técnicas sobre cómo modelizar con SIG variables físico-geográficas básicas, como la pendiente, y otras más complejas, como el tránsito potencial por el paisaje.

Además de identificar las propiedades de primer orden, es frecuente estudiar los factores de segundo orden, aquellos que se ocupan de describir la intensidad de la influencia ejercida por un punto sobre otro, analizando, de esta forma, los denominados efectos de segundo orden (O'Sullivan, Unwin 2003).

El análisis de los efectos de segundo orden permite valorar en qué medida existe una dependencia espacial entre los puntos, fenómenos como la atracción o repulsión entre ellos (Nakoinz, Knitter 2016). En casuísticas arqueológicas, los efectos de segundo orden pueden ser interpretados, por ejemplo, como nociones de territorialidad en el caso de asentamientos (Spencer, Bevan 2018), o fenómenos de nuclearización en torno, por ejemplo, a sitios megalíticos (Carrero-Pazos *et al.* 2019).

3.3.1. Un ejemplo práctico no arqueológico

Los efectos de primer y segundo orden sobre una distribución de puntos pueden entenderse mejor mediante un ejemplo específico. A continuación, los explicaremos sobre un caso de estudio que nada tiene que ver con Arqueología, pero que es perfectamente válido para entender los diferentes entresijos que atañen a cualquier análisis de patrones de puntos, a partir de la distribución espacial mostrada por los asistentes a un concierto de música al aire libre **(Figura 24**).

Las descripciones iniciales deben llevarnos a considerar el patrón de puntos que observamos en la imagen. Se trata de responder a preguntas del tipo ¿Por qué existe gente agrupada al norte de la imagen, o al sur? ¿Qué es lo que podría estar provocando esto? ¿Por qué existen zonas en las que la densidad de gente es mayor, y en otras es menor? ¿Podemos decir que el patrón que muestra la totalidad de la gente es agrupado? ¿O existen diversos patrones dependiendo de nuestra ventana de análisis –*zoom*-?

Continuando con las explicaciones anteriores, recordemos que todo patrón de puntos es el resultado de una combinación de dos tipos de factores. Factores exógenos o inducidos, externos a la distribución de puntos, y aquellos endógenos o inherentes a la distribución.

Las propiedades de primer orden, o la dependencia espacial inducida, se muestran en el caso del concierto a través de una variación de la densidad de gente en ciertas zonas del perímetro. Convendrán los lectores que la densidad de gente no es la misma en toda la superficie del

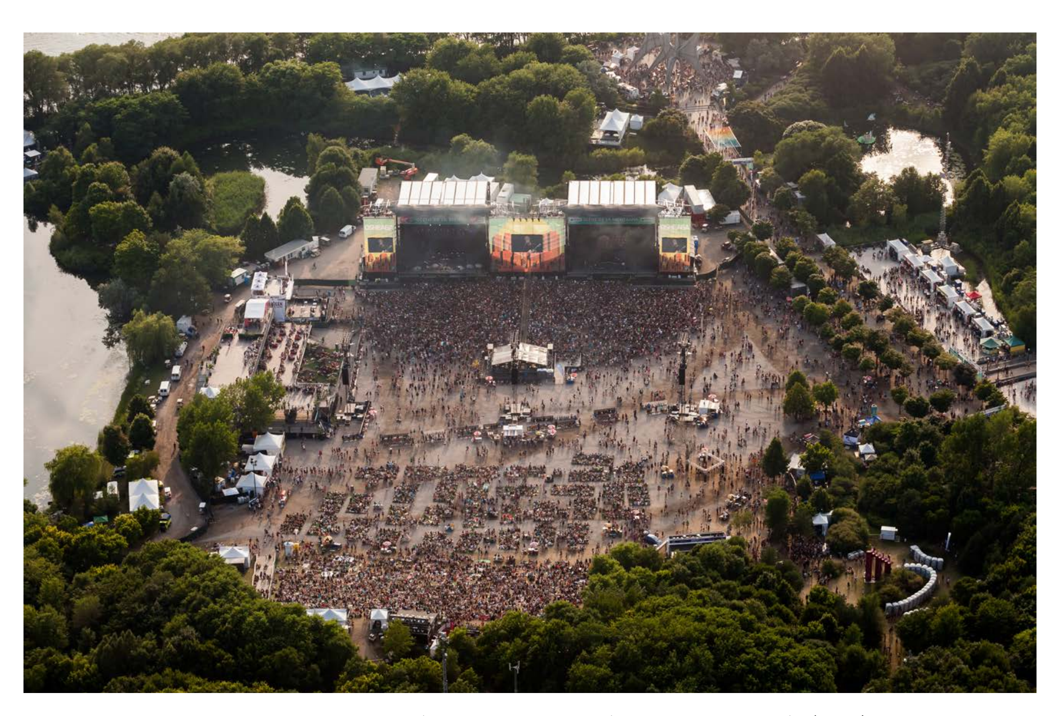


Figura 24. Vista aérea del *Osheaga Festival* de Montreal, Canadá (2006) (licencia standard, ©fotoimagemtl/123RF.COM).

concierto. Por ejemplo, existen áreas, como el *backstage*, en las que el acceso está restringido, por lo que estas zonas, al norte en la fotografía, tienen menos gente. El *main stage*, por su parte, es un factor de primer orden que influye en la composición del patrón de puntos, puesto que está forzando a la gente a disponerse a su alrededor. Las zonas de *impasse* visibles entre la masa de gente frente al *main stage* también podrían considerarse como variables de primer orden, puesto que su acceso está prohibido; lo que se entiende, por tanto, como una relación negativa entre estas áreas y la gente. Los factores de primer orden suelen actuar en escalas espaciales amplias, porque son factores que influyen en la totalidad de la distribución analizada, aunque a menudo lo hagan de forma desigual.

Las propiedades de segundo orden, recordemos, son aquellas que se constituyen a partir de la dependencia espacial entre los propios puntos. Es decir, en el caso del concierto, la localización individual de una persona vendrá condicionada por la existencia de otras personas (grupos de amigos). De ahí que existan agrupaciones pequeñas de gente en algunas áreas y, en otras, agrupaciones mayores.

La dependencia espacial puede adoptar dos tipologías espaciales básicas: la repulsión, como sería el resultado, por ejemplo, del protocolo covid-19 de distanciamiento social, o la atracción (agregación), resultado de la integración social. Los efectos de segundo orden suelen detectarse a pequeña escala y pueden mostrar diferente intensidad.

De la imagen del concierto se deducen, también, otros dos aspectos que vale la pena considerar. En primer lugar, que un patrón de puntos debe ser entendido siempre desde un punto de vista "multiescalar". Esto significa que los resultados de nuestros análisis dependen de la ventana analítica que utilicemos (en este caso, la porción de imagen que usemos para calcular el patrón -en términos simples, el *zoom* de nuestro análisis-). Si consideramos sólo la parte norte de la imagen, el terreno cercano al *main stage*, el patrón que muestra la distribución de gente será, con seguridad, agrupado. Pero si escogemos, por ejemplo, la zona central, el patrón podría tornarse regular. Y si nuestra ventana analítica es más pequeña, incluso podría ser aleatorio. Por otra parte, al sur de la fotografía aérea puede verse un conjunto numeroso de personas agrupadas en lo que parecen ser estructuras rectangulares (probablemente zona de mesas para comer). Aquí podemos estar ante un patrón agrupado si en nuestro análisis consideramos simplemente uno de estos rectángulos, o regular, si los consideramos todos, porque están situados a distancias muy similares unos de otros. Noten los lectores, entonces, el segundo aspecto en el que queremos incidir: los patrones de puntos no suelen ser homogéneos, existiendo disparidades importantes entre diferentes áreas de una misma zona de estudio.

3.3.2. Métodos de análisis

La metodología más utilizada para evaluar y estudiar las propiedades de primer orden es el modelado predictivo o la regresión logística (Mehrer, Wescott 2006; Verhagen, Whitley 2011), a partir de un trabajo cuantitativo con Sistemas de Información Geográfica del que hablaremos más adelante. El modelado predictivo suele orientarse, por ejemplo, a desentrañar correlaciones entre la probabilidad de descubrir determinados sitios arqueológicos y la distribución de una o más variables medioambientales, como pueden ser el tipo de suelo, la pendiente o la accesibilidad a recursos acuíferos, etc. (Bevan *et al.* 2013: 30). En segundo lugar,

es frecuente estudiar la distancia entre los puntos, es decir, la relación entre cada punto con los de su entorno.

A la hora de analizar las relaciones o interacciones entre un conjunto de puntos es necesario realizar aproximaciones que estudien la distancia entre ellos, midiendo entonces la importancia de las características de segundo orden, que describen la intensidad relativa de puntos influenciados por la configuración espacial de otros puntos en el área de estudio (Bevan *et al.* 2013: 31), y los efectos de segundo orden, que valoran, como ya hemos dicho, en qué medida existen fenómenos de atracción o repulsión entre ellos. Los efectos de segundo orden son el resultado de la interacción local entre las observaciones (O'Sullivan, Unwin 2003: 36), o lo que es lo mismo, la existencia de un evento puede provocar el crecimiento o la disminución de los eventos cercanos.

Para el estudio de las propiedades de segundo orden se utilizan habitualmente aproximaciones que analizan el espacio entre puntos. Es necesario indicar que las propiedades de segundo orden, y sus efectos, son más difíciles de detectar que las de primer orden, sobre todo si no se han tenido en cuenta correctamente primero los efectos de primer orden (Bevan 2020).

De entre los métodos más comunes que estudian la estructura de un patrón de puntos destaca el análisis de cuadrantes, que define una zona de estudio en unidades regulares y estudia el número de puntos que aparecen dentro de ellas, si bien cuenta con algunos problemas como los efectos de la escala de análisis (Olaya 2014: 295). El cálculo se concreta en contar cuántos eventos caen dentro de cada cuadrante y comparar la frecuencia de la distribución con aquella que cabría esperar si los puntos estuviesen situados aleatoriamente. En términos simples, se coloca una cuadrícula a través del patrón de puntos y se utiliza una prueba de chi-cuadrado para evaluar el grado en que los recuentos de la cuadrícula difieren de lo esperado bajo condiciones aleatorias (**Figura 25**).

Cabe indicar ahora que el análisis de cuadrantes no es en realidad una medida del patrón, sino de la dispersión. Además, debido al uso de una unidad fija (el cuadrante), el análisis puede no ser capaz de localizar agrupamientos locales. Otra debilidad del método es que los resultados se ven seriamente afectados por el tamaño de la ventana que se utilice (Hodder, Orton 1976), siendo común el uso de un tamaño del doble del área media disponible para cada punto, es decir, cuadrados donde cada lado poseerá una longitud $l = \sqrt{2A/N}$, donde N es el número de puntos y A el área de la zona de estudio (Olaya 2014: 295).

En la **Figura 25** puede observarse el resultado gráfico de un cálculo de cuadrantes sobre un mapa de densidades *Kernel*. En el gráfico, el recuento observado de cada cuadrícula se muestra en la parte superior izquierda; el esperado en la parte superior derecha, y los valores residuales en la parte inferior, que indican la diferencia entre los valores observados y los esperados. De esta forma, los residuales tienen una media de 0 y una varianza de 1, lo que significa que las desviaciones importantes son las que tienen lugar entre -1 a 1 en cualquier dirección, mientras que las variables dentro del rango de -1 a 1 son las esperadas. En la **Figura 25** observamos que el área de mayor densidad en el centro cuenta con desviaciones altas respecto a lo esperado, según los residuales. Podemos concluir, entonces, que los datos no son homogéneos, pues existe una densidad elevada en torno al centro.

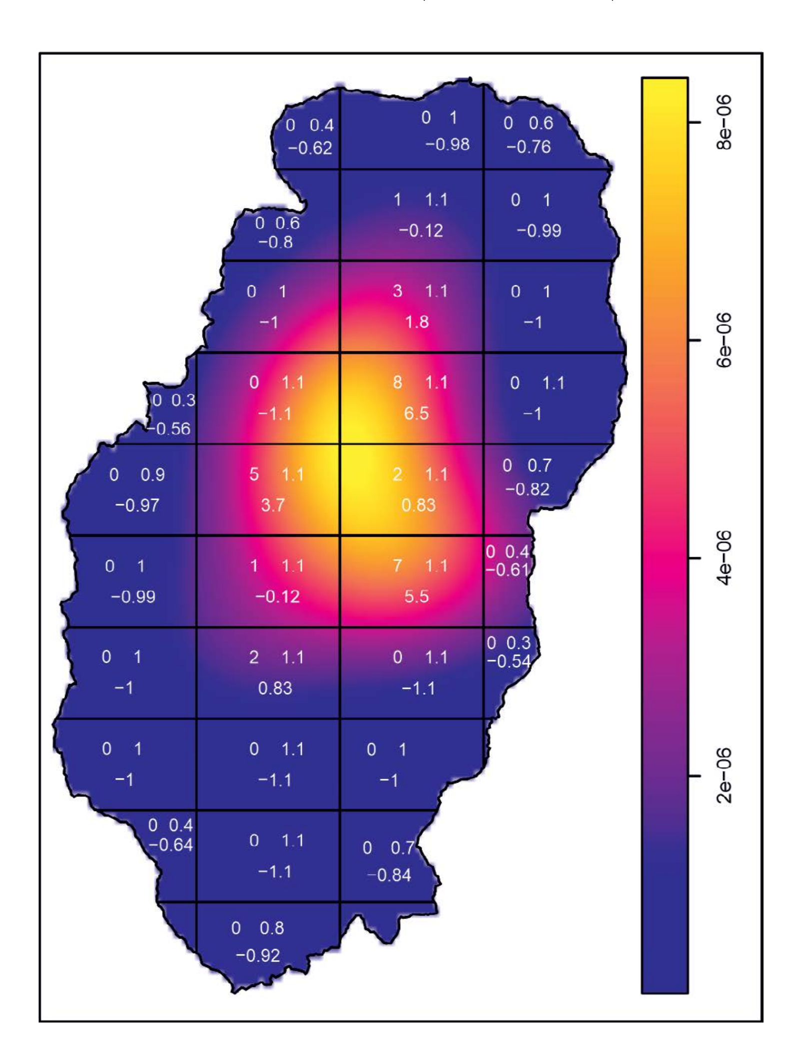


Figura 25. Ejemplo de análisis de cuadrantes.

En segundo lugar, el análisis de vecino más próximo ha sido utilizado tradicionalmente en Arqueología para detectar agrupaciones o caracterizar distribuciones de yacimientos (cfr. Hodder, Hassall 1971; Hodder 1972; Whallon 1974; Peebles 1973, entre otros). Este análisis considera las distancias entre cada punto y su vecino más próximo, calculando, posteriormente, el promedio de esos valores (Clark, Evans 1954) y comparando cada medida con aquella que cabría esperar para el conjunto de la región estudiada (**Figura 26**).

El test de Clark y Evans es uno de los procedimientos matemáticos más utilizados para calcular el vecino más próximo, pues proporciona una medida que busca sugerir si la distancia más próxima es inferior a la que podríamos esperar (es decir, los sitios se encuentran agrupados), mayor de lo que cabría esperar (los sitios están regularmente espaciados) o no mantiene ninguna de estas características (se encuentran aleatoriamente espaciados). Según A. Bevan y J. Conolly (2006: 218), este estadístico es el más utilizado porque es de sencilla aplicación y proporciona resultados fácilmente interpretables.[33] No obstante, un problema importante con

[33] En *R statistics* su cálculo es muy sencillo, a partir del *clarkevans.test*, que nos proporciona un valor de distancias de *R*

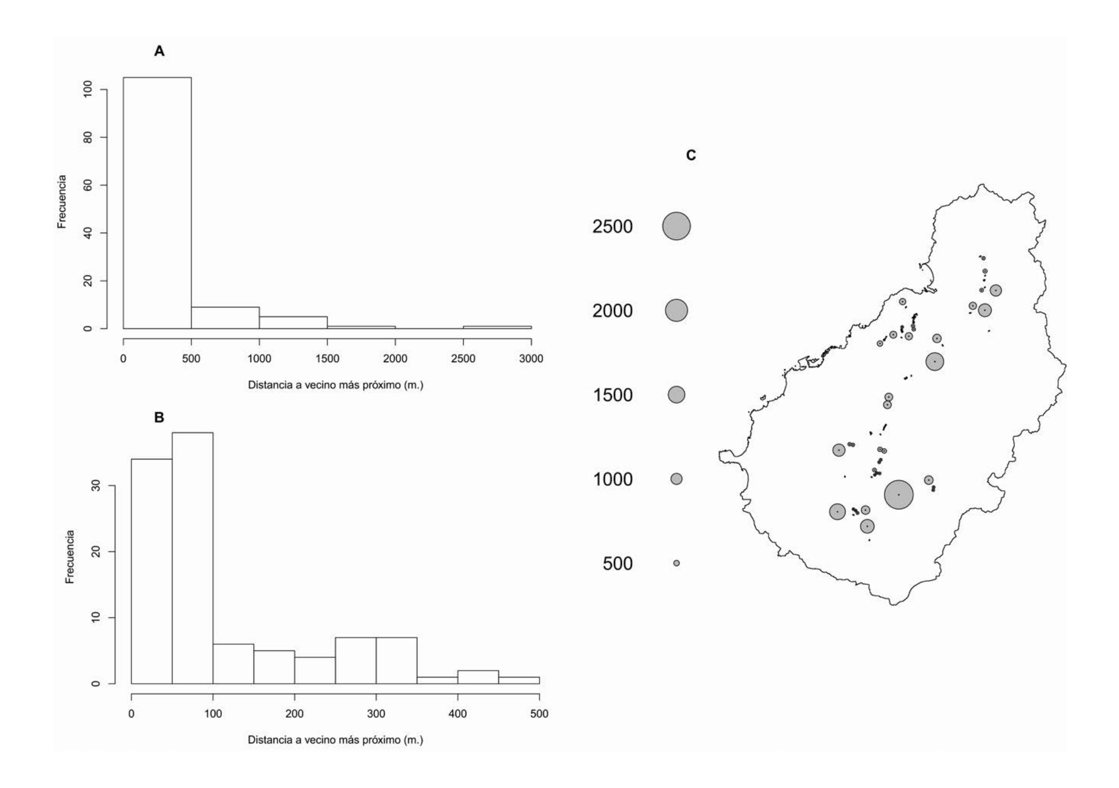


Figura 26. Ejemplo de análisis del vecino más próximo. **A:** Histograma de la medida general, en el que se observa que la mayor parte de los puntos se localizan a menos de 500 m unos de otros. **B:** Histograma centrado en los primeros 500 m, para observar con más detalle dicha tendencia. **C:** Diagrama de Stienen, que refleja la distribución del patrón espacialmente. Los círculos de mayor tamaño indican mayor distancia a sus vecinos, mientras aquellos más pequeños se localizan más cerca de sus vecinos. (Carrero-Pazos 2017: 199).

respecto a este test es que los resultados pueden estar sesgados por los efectos de borde, ya que es frecuente que el vecino más cercano de un punto se sitúe cerca del extremo del área de estudio, y por lo tanto puede ocurrir que algunos caigan fuera del mismo.[34]

Por otra parte, es necesario tener en cuenta la influencia del tamaño del área de estudio en la detección y caracterización de patrones. Si observamos la **Figura 27**, veremos cómo el ajuste de la escala de análisis condiciona directamente la homogeneidad, intensidad y tendencias de agrupación de las distribuciones de puntos (Bevan, Conolly 2006: 220).

(índice de vecino más próximo o NNI) y un valor *p*. Este valor *p* es un índice de significación, que nos indica agrupación cuando R<1, distribución regular cuando R>1 y aleatoria cuando R±1 (Bevan 2020). No obstante, aunque el valor *R* sugiera agrupación, si el valor *p* que lo acompaña sugiere que el patrón no es significativo, entonces no podemos indicar que nuestro valor de vecino más próximo difiera de lo que cabría esperar en condiciones azarosas.

[34] Esto significa que las distancias de vecino más próximo observadas son probablemente más grandes que las distancias reales, o lo que es lo mismo, la presencia de un límite que no permite hacer mediciones más allá de él hará que el valor de NNI tienda a ser mayor. Existen correcciones aplicables que intentan solucionar el problema del efecto de borde, como la presentada por K. Donnelly (1978), aunque ninguna proporciona resultados finales (Bevan *et al.* 2013).

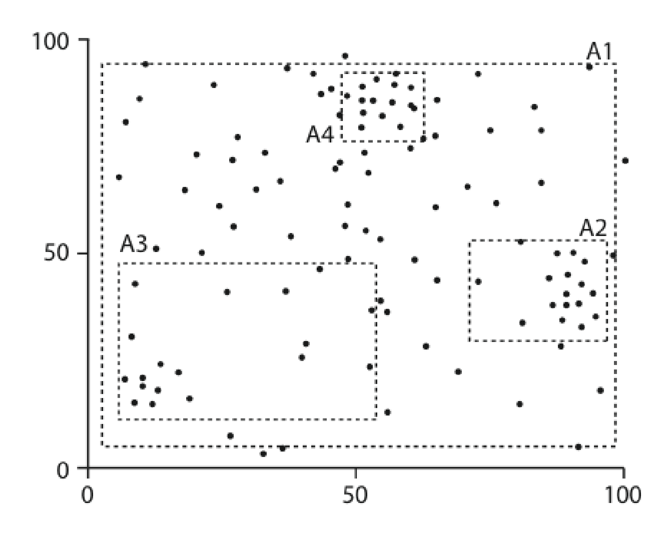


Figura 27. Influencia del tamaño del área de estudio en la detección de patrones de puntos (Goreaud, Pélissier 2000, tomada de Bevan, Conolly 2006: 220).

En la **Figura 27** (A1), la agrupación tiene una estructura homogénea a lo largo de toda el área de estudio, con zonas que presentan mayor intensidad de puntos que otras. El resultado de un test de vecino más próximo indicaría que la distribución de puntos es regular. Si atendemos al sector A2, el patrón ya no es homogéneo sino heterogéneo, con una intensidad fuerte en su mitad derecha. Los resultados ya no serán similares a los del sector A1.

Finalmente, y pese a que estamos ante un análisis ampliamente utilizado en Arqueología, quizás el inconveniente más importante del índice de vecino más próximo de Clark y Evans es que nos proporciona un solo índice sobre la distancia entre los vecinos más cercanos para todo el conjunto de datos, y no tiene en cuenta la escala del proceso ni que las distribuciones de puntos, al menos en Arqueología, nunca son homogéneas (Bevan 2020). Puede ocurrir, por ejemplo, que una agrupación aislada de ciertos asentamientos muestre una tendencia regular desde una perspectiva general, aunque esconda agregación si modificamos la escala (**Figura 28**).

Un segundo conjunto de análisis que se ocupa del estudio de las distancias entre puntos son las funciones K y sus derivadas, que se caracterizan por permitir el examen de procesos de procesos de punto a múltiples escalas (Orton 2004: 303; Baddeley *et al.* 2015), razón por la cual muchos autores utilizan estas técnicas frente al estadístico de vecino más próximo (por ejemplo, Bevan, Conolly 2004, 2006, 2009; Sayer, Wienhold 2012; Bevan *et al.* 2013).

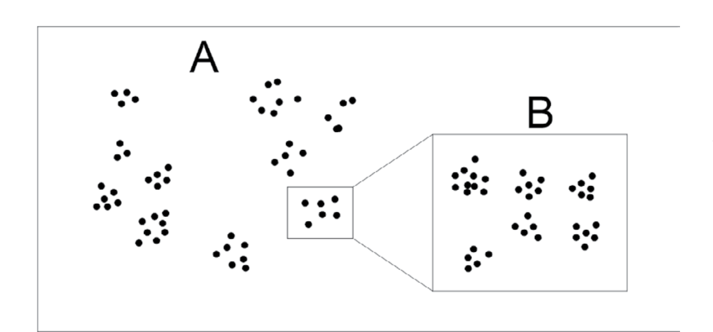


Figura 28. Patrones de puntos "multiescalares". Patrón general (**A**), que esconde agrupación si reducimos la escala de análisis (**B**).

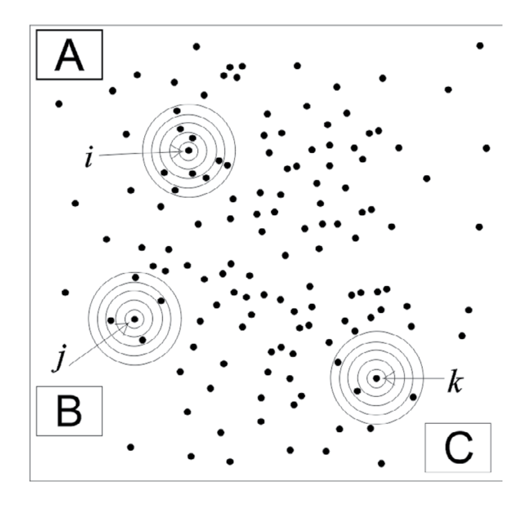


Figura 29. Estimación de la función K de Ripley.

Dentro de las funciones de Ripley, destaca la función K (Ripley 1976, 1977, 1981), que, en vez de fijar una escala de análisis y una serie fija de cuadrantes, se concreta en una serie de zonas de análisis que son estudiadas a diferentes escalas y con distintos tamaños. La función K computa la media esperada del número de eventos que caen dentro de una determinada distancia, *d*, para cada evento, escalando el cálculo por la densidad de los eventos (**Figura 29**).

Se define, matemáticamente, por un proceso de intensidad λ, donde λK(r) es el número esperado de vecinos en un círculo con radio *r* en un punto arbitrario de la distribución (Pélissier, Goreaud 2001: 101). Es, por tanto, una distribución de frecuencias acumuladas del promedio de la intensidad de puntos en una serie de intervalos de *r*.

Los intervalos de significación son generados por distribuciones aleatorias de puntos gracias a la Simulación de Monte Carlo. Así, estas estimaciones pueden ser comparadas con los valores K observados para proporcionar una medida del tamaño y distancia de las agrupaciones de la muestra de forma más robusta[35] (**Figura 30**).

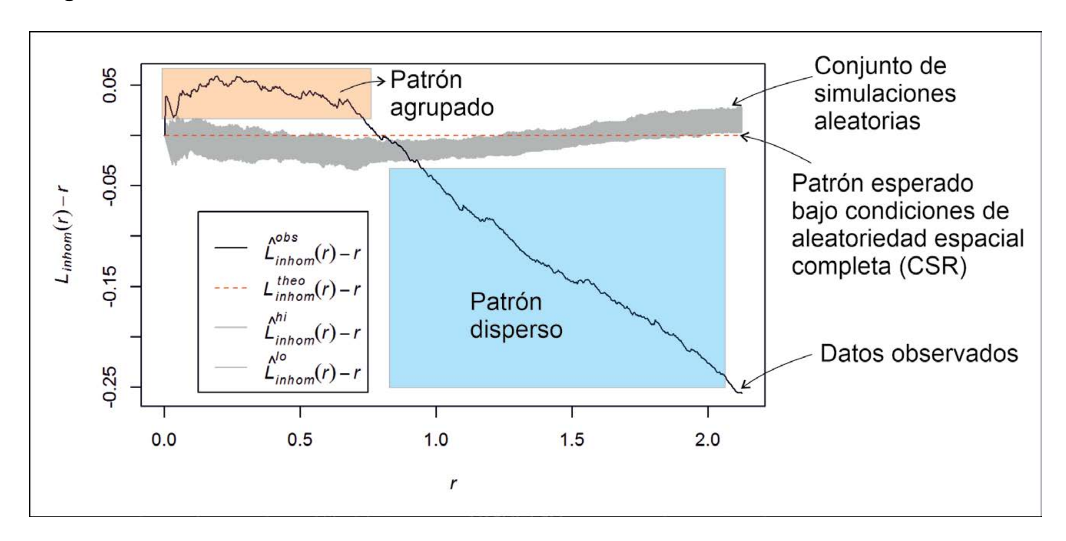


Figura 30. Gráfico de función L de Ripley (variante de la K), con explicación de sus diferentes parámetros. Modificado a partir de un gráfico de M. Peeples (mattpeeples.net).

[35] Por norma general, se utiliza la corrección del efecto de borde propuesta por F. Goreaud y R. Pélissier (2000), con la transformación de la distribución acumulada *K* a L(r)=√(K(r)/π-r), en la que el valor esperado bajo condiciones aleatorias [L(r)=0] es una línea horizontal. Si L(r)>0, existen más eventos a una distancia determinada *d* que los

Asimismo, la Función K de Ripley y sus variantes permiten comparar la tendencia de los datos observados con un conjunto de simulaciones aleatorias, es decir, el patrón esperado bajo condiciones de aleatoriedad espacial completa (CSR). En la **Figura 31**, que representa un gráfico de la función L de Ripley, observamos cómo la línea negra (distribución observada, por ejemplo, asentamientos neolíticos), se sitúa, en ciertas partes, por encima del conjunto de simulaciones aleatorias (*random envelope*, en la terminología inglesa más frecuente), y en otras por debajo. Cuando está por encima (valores *r* de 0 a ca. 0,7 en el gráfico propuesto), podemos concluir que la tendencia de los datos observados es agrupada (Bevan *et al.* 2013; Bevan 2020). Cuando está por debajo, la tendencia mostrará dispersión o regularidad (distancias de aproximadamente 0,9 a 2 en este ejemplo). Los dos patrones que pueden identificarse, tanto el agrupado como el disperso, sabemos que son estadísticamente significativos porque la línea negra no cae, en estos intervalos de distancias, sobre el conjunto de simulaciones aleatorias. Cuando esto ocurre, caso del intervalo de *r* 0,8-0,9 (la parte de la línea negra que se sitúa dentro del rango gris), entonces el patrón observado pudo haber sido formado bajo condiciones de aleatoriedad, por lo que, en términos simples, no es significativo.

El mismo ejemplo nos permite profundizar en la correcta interpretación de estas funciones. En la **Figura 31** puede observarse una función L con datos observados para la distribución de cierto tipo de árboles japoneses, mostrando un patrón significativo de dispersión o regularidad a escalas locales, frente a un modelo aleatorio. Esto viene a indicar que, en este caso particular, los árboles mantienen un patrón de inhibición entre ellos a escala local, que provoca que estén situados a distancias similares unos de otros. Y dicha tendencia no es el resultado de un proceso aleatorio, con lo que habrá que seguir investigando qué puede estar provocando este patrón. Por ejemplo, sería muy recomendable analizar, con mayor detalle, posibles efectos de primer orden que puedan estar influyendo a escalas locales.

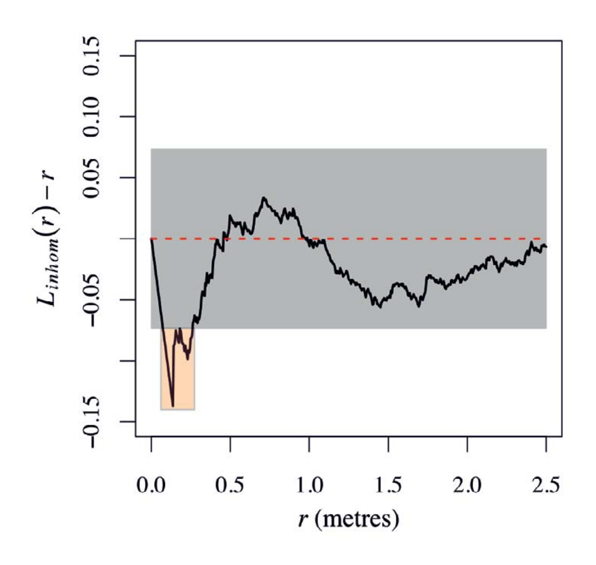


Figura 31. Ejemplo de una función L de Ripley, para el estudio de la distribución de cierto tipo de pinos japoneses, junto con 19 simulaciones aleatorias creadas bajo un proceso *Poisson* (a partir de Baddeley *et al.* 2015: 246). En naranja puede identificarse un patrón significativo de dispersión o regularidad a escala local.

esperados bajo condiciones de aleatoriedad, lo que significa que existe un patrón agrupado; Si L(r)=0, existen tantos eventos a una distancia *d* como los esperados bajo condiciones de aleatoriedad, lo que indica un patrón aleatorio; Si L(r)<0, existen menos eventos a una distancia *d* que los esperados bajo condiciones de aleatoriedad, lo que sugiere un patrón regular (Bevan, Conolly 2006: 221; Pélissier, Goreaud 2001: 102).

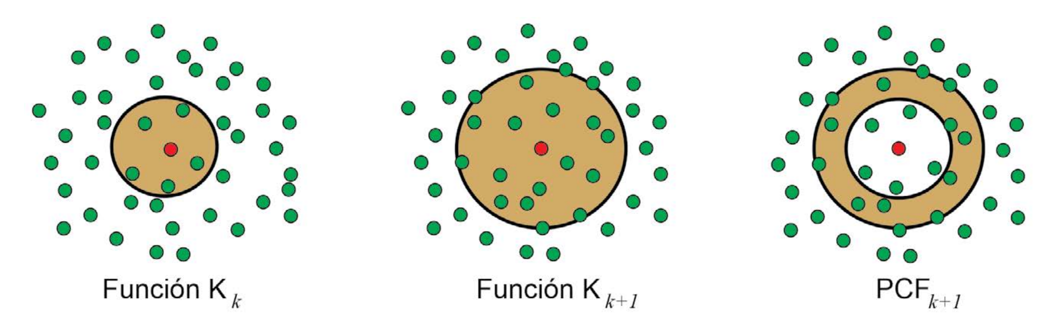


Figura 32. Representación esquemática del cálculo de las funciones K y de correlación par (*pcf*).

Una variante del conjunto de métodos de Ripley que está siendo muy utilizada en trabajos arqueológicos (por ejemplo, Bevan *et al.* 2013; Spencer, Bevan 2018; Carrero-Pazos *et al.* 2019; Bevan 2020; Riris 2020) es la función de correlación par (*pair correlation function,* frecuentemente por sus siglas *pcf*), que difiere de la K en que mide la asociación espacial dentro de anillos (**Figura 32**).

La función de correlación par permite estudiar un patrón de puntos a múltiples distancias mediante el uso de frecuencias de puntos localizadas dentro de bandas sucesivamente más grandes alrededor de cada punto. Con ello, busca resumir la probabilidad de que los sitios arqueológicos estén agrupados o dispersos a diferentes escalas espaciales en función del número de puntos localizados dentro de cada distancia determinada.

Las diferentes escalas espaciales se refieren, potencialmente, a diferentes tipos de interacción entre sitios arqueológicos, desde por ejemplo relaciones de comunidad muy localizadas (como las que se identificarían en los análisis tradicionales del vecino más cercano) hasta patrones más amplios y regionales, por lo que pueden poner de manifiesto diferentes procesos económicos, sociales o políticos al mismo tiempo (Spencer, Bevan 2018: 76-77). En la **Figura 33** podemos encontrar un gráfico de correlación par junto a la explicación de sus parámetros principales.

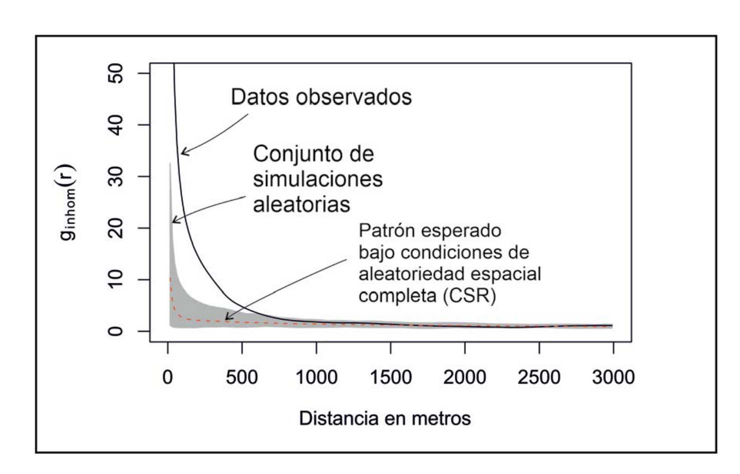


Figura 33. Función G de correlación par, junto a la identificación de sus componentes analíticos principales.

Su explicación es similar al de la función K. La línea negra identifica los datos observados (por ejemplo, sitios arqueológicos), mientras que el rango gris se compone de un conjunto de simulaciones aleatorias creadas por técnicas de Monte Carlo. Si la tendencia de los datos observados (línea negra), se sitúa por encima del conjunto de simulaciones aleatorias (en gris), concluiremos entonces que el patrón de la distribución estudiada es agrupado. Si está por debajo, será disperso (Bevan *et al.* 2013; Bevan 2020). En este caso concreto, los datos observados muestran una distribución agrupada hasta los 550 metros, aproximadamente, cuando pasan a poseer una tendencia similar a la del patrón aleatorio. Note el lector, también, cómo dicha tendencia cambia a medida que analizamos otras distancias, reflejando fenómenos de agrupación hasta los 1.500 metros, dispersión entre 1.500-2.000 metros y, de nuevo, agrupación. Aunque se concluye que estos patrones no son significativos, puesto que existe probabilidad de que pudiesen haber sido creados en condiciones aleatorias.

Todas las funciones comentadas en este capítulo y, en términos generales, los análisis de patrones de puntos pueden calcularse en *R Statistics* utilizando el paquete *spatstat* (Baddeley, Turner 2005; Baddeley *et al.* 2015).

Otro método que es necesario tener en cuenta es la autocorrelación espacial, que parte directamente de los enunciados de Tobler en 1970, y considera que los valores de atributo de una localización tienden a ser más similares con los de una localización cercana que con valores de atributos de localizaciones a distancias lejanas. En pocas palabras, los sitios cercanos suelen parecerse más entre ellos que con respecto a los sitios que se encuentran alejados, debido a la autocorrelación espacial.

La autocorrelación espacial puede ser positiva, cuando los valores cercanos son similares, o negativa o cero, cuando los valores cercanos son aleatorios. En caso de no haber autocorrelación, los datos son independientes entre sí y no se afectan mutuamente. Las medidas de autocorrelación espacial más utilizadas son el parámetro I global de Moran y el parámetro *c* de Geary (1954). El uso de estos análisis en Arqueología no está muy extendido, dado que ambos parámetros caracterizan la autocorrelación para todo el conjunto del área de estudio y proporcionan un resultado general, y ya hemos dicho que tendencias como esta son infrecuentes a la hora de estudiar comportamientos humanos del pasado. Existen otros dos interesantes que miden la autocorrelación a escala local, y son los de Getis y Ord (1992), que cuantifican si un punto se encuentra rodeado por agrupaciones de puntos con valores altos o bajos, o una versión local del parámetro I de Moran, denominado "indicador local de asociación espacial" (*Local Indicator of Spatial Association*, LISA) (Anselin, 1995) (para una aplicación arqueológica, consúltese, por ejemplo, Premo 2004).

Finalmente, directamente relacionado con el anterior tenemos el análisis de densidades *Kernel*, que se concreta en la estimación de una medida de densidad de puntos mediante el movimiento de una ventana *Kernel* (de diferente tipo) con un radio *r,* por todo el mapa raster (Silverman 1986). En cada localización, la ventana analítica cuenta el número de eventos que caen dentro del *Kernel* y lo divide por el tamaño total del área, consiguiendo, de esta forma, una estimación de aquellas áreas con mayor densidad de puntos y aquellas áreas con menor densidad de puntos (**Figura 34**).

Kernel density surface of sites (sigma=2000m)

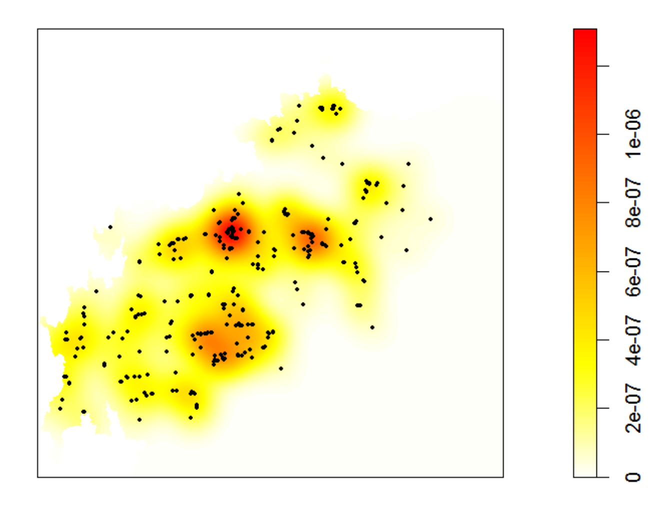


Figura 34. Estimación de densidad del núcleo (*Kernel Density Estimation*, KDE) para los monumentos megalíticos localizados en A Costa da Morte (Galicia), utilizando un sigma (*r*) de 2 km.

3.3.3. Principales problemas en el análisis espacial

El primero viene derivado del principio de autocorrelación espacial, que reduce la independencia entre las observaciones, por lo que añadir nuevas observaciones próximas en distancia a nuestro conjunto de datos puede no proporcionar nueva información sustancial. En segundo lugar, y derivado de la definición de las unidades de análisis, el Problema de la Unidad de Área Modificable (PAUM), que es causado por el uso de unidades artificiales en la definición espacial de fenómenos geográficos continuos que, como consecuencia, pueden generar patrones artificiales (Heywood 1998; Olaya 2014).[36]

Finalmente, un último aspecto a reiterar son los efectos de borde. Cualquier zona que estudiemos dentro de un análisis espacial tiene siempre unos límites establecidos que, en algunos casos, serán naturales (elementos geográficos, cuencas hidrográficas) y en otros artificiales (divisiones administrativas actuales, por ejemplo). Estos límites pueden crear falsos patrones si existen puntos cerca de ellos, aunque existen soluciones estadísticas para tenerlos en cuenta (**Figura 35**).

[36] El PAUM es un problema estudiado desde hace tiempo, y a pesar de ser extremadamente relevante a la hora de realizar estudios espaciales, las soluciones no son claras. Tradicionalmente se considera que se trata de un problema inasumible (Olaya 2014: 245). Por otro lado, existen además dos cuestiones particulares que están relacionados con el PAUM, denominadas *falacias*. Por un lado, la falacia ecológica (Openshaw 1983) consiste en asumir que la relación observada a un nivel de agregación está presente también en niveles inferiores de agregación, y, por otro lado, la contextual, que supone aceptar que los valores calculados para una unidad de área pueden aplicarse a los individuos de la población existente en dicha área, es decir que los individuos siempre se comportan de la misma forma independientemente del lugar en el que se encuentren.

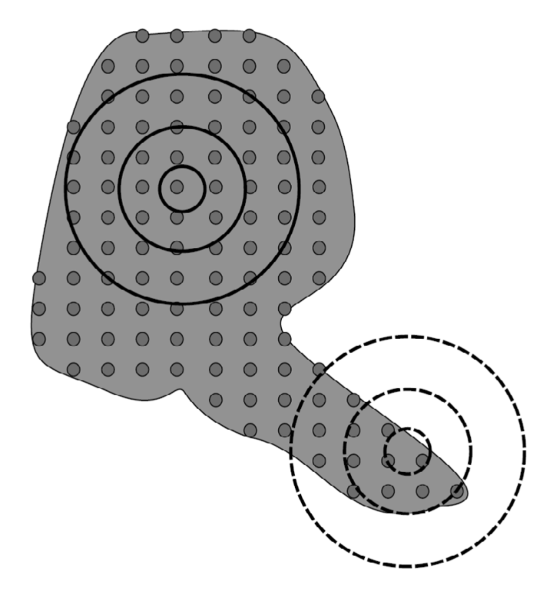


Figura 35. Problemas derivados del efecto de borde, y cómo este afecta en mayor o menor medida en función de la escala de análisis (tamaño del círculo) (modificado a partir de Olaya 2014: 249).

3.4. El análisis locacional

Una de las ramas de la investigación arqueológica más prolija en la actualidad es el "análisis locacional", que surgió tras el nacimiento de la Nueva Arqueología en los años sesenta del siglo pasado. El término procede del inglés *locational analysis*, comúnmente aceptado y utilizado por la bibliografía especializada. El análisis locacional es parte contante de uno de los cambios más importantes de la disciplina arqueológica del último siglo, como fue desarrollar la Arqueología espacial (Hodder, Orton 1976; Clarke 1977).

En la mayor parte de los trabajos iniciales en Arqueología espacial se observa un marcado carácter economicista, que no sorprende a tenor del desarrollo e importancia que adquirieron aportaciones como el *Site Catchment Analysis* desarrollado por V. Flannery y M. D. Coe (1969), a la par que C. Vita-Finzi y E. Higgs (1970). Desde esta óptica, los sitios arqueológicos, como por ejemplo asentamientos, son estudiados en torno a conceptos como el elemento central en la explotación de ciertos recursos económicos. No obstante, el análisis locacional varió y evolucionó a lo largo de las siguientes décadas, adecuándose paulatinamente a los cambios epistemológicos de la Arqueología y aprovechando las oportunidades abiertas por las nuevas Tecnologías de Información Geográfica (TIG) (García García 2015: 93). El espacio empezaba a ser entendido no sólo como condicionante de la acción social, sino como su continente, resultado de las actividades humanas.[37]

Todo análisis locacional de un sitio o conjunto arqueológico parte siempre de un modelo teórico de asentamiento, poblamiento o localización específico. En su primer término se basa, pues, en un importante trabajo bibliográfico, con la creación de una síntesis de estudio de la que posteriormente se extraerán diferentes variables analíticas.[38] Por ejemplo, en

[37] Todo este cambio teórico se desarrolló a la par que una auténtica revolución tecnológica, cuando comenzaba lo que Kvamme denominó *the second age of modelling* (Kvamme 2006: 4), con el desarrollo de las metodologías geoespaciales (SIG, GPS, Teledetección), que facilitaron la aplicación del análisis locacional de forma más precisa.

[38] Las orientaciones metodológicas seguidas en esta parte se han desarrollado principalmente a partir de los trabajos

el caso del Megalitismo del Noroeste peninsular, tema que ocupó nuestra tesis doctoral, contamos con reflexiones sobre estos paisajes funerarios por parte de los pioneros de la arqueología científica gallega de mediados del siglo pasado, nombres como F. López Cuevillas o F. Maciñeira, entre muchos otros. Posteriormente debemos acudir a trabajos más modernos en los que se analiza, a través del trabajo de campo, la importancia de variables como la visibilidad o la proximidad a rutas de tránsito en la localización de los megalitos.

En este contexto, el análisis locacional se estructura como un conjunto de procedimientos analíticos orientados a rastrear las decisiones sociales y culturales que permiten entender el proceso de selección de una determinada posición espacial para la construcción de, por ejemplo, un túmulo megalítico, castro o abrigo paleolítico.

El análisis locacional debe ser planteado como una metodología "arqueo-geográfica" (Vicent García 1991), en la que, gracias a las tecnologías geoespaciales, se analiza la localización de determinados sitios arqueológicos y las relaciones que se derivan de ellos con otros elementos del paisaje, con el objetivo de testar las hipótesis acerca de las decisiones que se tomaron a la hora de construirlos (García García 2015); o lo que es lo mismo, analizar la dimensión espacial de la acción social (Llobera 2011). El objetivo fundamental del análisis locacional reside, pues, en intentar proponer una respuesta coherente, a partir de criterios cuantificables, a la pregunta elemental de por qué ciertos sitios arqueológicos están donde están (Fonte 2015: 103).

El procedimiento analítico se concreta, desde una perspectiva teórica, en considerar que la localización de un sitio arqueológico determinado, sea un asentamiento, taller, o monumento funerario etc., estuvo condicionada por una serie de decisiones humanas que hicieron primar ciertos criterios espaciales sobre otros. Estos criterios pueden, de hecho, concretarse en una serie de factores del paisaje que son materiales y, por ello, cognoscibles arqueológicamente (Parcero Oubiña, Fábrega Álvarez 2006: 72).

En último término, esto supone aceptar que la localización de un yacimiento arqueológico viene dada por decisiones sociales y culturales, y su determinación un elemento relevante para la caracterización de una determinada formación social. Con el análisis locacional se pretende, por tanto, recomponer aquellas decisiones a partir de su materialidad en el registro arqueológico, en el caso más común, estudiando la distribución en el paisaje de cierto tipo de sitios arqueológicos.[39]

Cuando, por ejemplo, el objeto de estudio es un asentamiento, existen ciertos criterios que casi con total seguridad pudieron condicionar o constreñir las decisiones locacionales tomadas. La necesidad de agua, por ejemplo, suele ser uno de los más estudiados. De aquí se deduce que existen determinados factores que, a su vez, condicionan a estos criterios como, por ejemplo, la facilidad de acceder a un recurso acuífero o la disponibilidad inmediata del mismo. Entonces,

de C. Parcero Oubiña y P. Fábrega Álvarez (2006), y M. García García (2015).

[39] Esta metodología no es la única y existen variantes, como la presentada por P. Verhagen y T. Whitley (2011). Estos autores hacen una distinción entre "condiciones", aquellas características que pueden ser observadas en el paisaje; "eventos", comportamientos o acciones limitadas en espacio y tiempo que de forma general no pueden ser directamente observables hoy, pero quizá haya tenido consecuencias físicas (condiciones) parcialmente observables; y, por último, "decisiones", que se corresponden con las elecciones que tuvieron lugar en un momento determinado, que unen las observaciones de las condiciones y el comportamiento de los eventos.

a partir del examen locacional de los lugares es posible aproximarnos a los criterios sociales y culturales, a las decisiones tomadas por la sociedad para construir sus asentamientos o monumentos en una zona determinada. De los criterios pasamos así a los factores, y de éstos a las decisiones (Parcero Oubiña, Fábrega Álvarez 2006). De esta manera, el objetivo inmediato será explorar en qué medida se priorizan unos factores sobre otros, aunque debemos tener en cuenta que algunas condiciones geográficas variaron a lo largo del tiempo y habrá que seleccionar aquellas que puedan haber permanecido más inmóviles. Por todo ello, dice M. García García (2015: 95), la intención no debe ser reproducir el paisaje sino analizarlo, no inferir sino contrastar hipótesis. En este contexto, las opciones metodológicas proporcionadas por los SIG y la simulación estadística proporcionan un buen contrapunto, ya que permiten la contrastación de hipótesis o preguntas, en definitiva, aquel modelo interpretativo construido sobre el conocimiento previo, al poder hacer observables y cuantificables los efectos de la acción social en el paisaje.

Finalmente, es necesario realizar una puntualización importante sobre los análisis locacionales. No parece ser frecuente encontrar la realización de un trabajo crítico sobre los factores que puedan haber influido en la localización de un determinado conjunto de sitios arqueológicos. Es frecuente recurrir a trabajos anteriores para justificar el estudio de cierto criterio y, por ende, su materialización en toda una serie de factores que serán analizados con métodos SIG. Esto puede solucionarse si introducimos en el discurso metodológico locacional el modelado predictivo (Mehrer, Wescott 2006; Verhagen, Whitley 2011) como el método que permite estudiar y criticar, en profundidad, la relevancia locacional de las variables manejadas, observando cómo y en qué medida influyen, y si lo hacen realmente, en la distribución observada de los sitios arqueológicos en cuestión.

Otra técnica estadística que puede utilizarse es la simulación de Monte Carlo, que nos permite comparar la tendencia de los yacimientos con la de múltiples muestras aleatorias distribuidas en el área de trabajo, generando, con ello, un índice de significación en torno a la variable analizada.

De la aplicación de ambos métodos se puede inferir que, si cierta variable medioambiental influye en la distribución de ciertos yacimientos arqueológicos, presumiblemente pudo haber sido tenida en cuenta a la hora de tomar la decisión de localizar, en el lugar donde hoy los encontramos, los sitios arqueológicos en cuestión.

SEGUNDA PARTE:

MODELANDO DECISIONES HUMANAS EN PAISAJES PRETÉRITOS

Capítulo 4.

Aproximaciones con Sistemas de Información Geográfica y estadística espacial para el estudio de patrones geográficos

4.1. Propiedades y efectos de primer orden. Las variables medioambientales

Toda investigación espacial debe partir, como hemos señalado anteriormente, del estudio en profundidad de los trabajos previos sobre el tema específico, con el objetivo de definir un modelo locacional interpretativo de los sitios arqueológicos en cuestión. Como ejemplo metodológico expondremos a continuación, en unas breves notas, aquel que definimos en su momento para el Megalitismo gallego, como muestra de un tipo de modelo que podría seguirse (véase Carrero-Pazos 2017 para explicaciones más extensas).

Las investigaciones espaciales del Megalitismo gallego no son recientes, sino que encontramos referencias ya desde antiguo. Pese a que los túmulos megalíticos hayan desempeñado importantes funciones para los vivos desde el mismo momento en que fueron construidos (Fleming 1973; Delibes de Castro 1991), con significados que van más allá de la sepultura funeraria, su relevancia espacial puede rastrearse en períodos posteriores. M. Martinón-Torres (2001) da cuenta de este hecho, poniendo de relieve la presencia de megalitos como hitos territoriales en fuentes escritas desde tiempos medievales y modernos, aunque será durante el S. XIX y XX cuando los primeros trabajos científicos en Arqueología coadyuven la aparición de discursos específicos sobre los patrones de localización del fenómeno megalítico.

Tenemos, así, referencias espaciales directas, como el caso de M. Díaz Sanjurjo (1904), que relaciona la localización de los túmulos con la proximidad a caminos antiguos, aspecto también defendido años después por F. Maciñeira (1935), que indicaba que los túmulos se emplazaban flanqueando los caminos existentes cuando fueron construidos. Otros autores hacen hincapié, por ejemplo, en patrones topográficos, como L. Martínez de Padín (1849), para quien los túmulos megalíticos se hallaban, con mayor frecuencia, en el fondo de los valles, llanuras o sobre las montañas.

Ya en la década de 1980, las investigaciones de J. M. Bello Diéguez, J. M. Vázquez Varela o F. Criado Boado, entre otros, señalan la importancia, por ejemplo, de los usos del suelo como

factor de localización. Nos dicen que los túmulos "en vez de asentarse en terrenos pobres y yermos, lo hacen en las proximidades de las mejores tierras de las zonas en las que se encuentran (Bello Diéguez *et al.* 1982; Criado Boado 1988: 662). Pero también encontramos referencias a factores físico-geográficos, como la relación de túmulos con las mayores altitudes absolutas, su localización específica en topografías prominentes o en zonas de visibilidad amplia (Criado Boado 1984: 14, 1988: 669). Otros estudios se ocupan, también, de la relación del emplazamiento de estos monumentos con rutas de tránsito por el paisaje a diferentes escalas (por ejemplo, Vaquero Lastres 1991-1992).

Estos son, entre otros, diversos criterios de localización que pueden encontrarse en la bibliografía y que fueron desarrollados desde investigaciones basadas en trabajo de campo. M. Llobera (2006: 110-111) resume de forma acertada la tendencia general de estos trabajos pues "a pesar de su innegable uso, muchos de estos intentos son, en el mejor de los casos, descripciones con un alto nivel de detalle y profundidad que más que resolver sirven para formular preguntas con gran significancia y atino. En el peor de los casos, se utilizan para generar interpretaciones nacidas más de la especulación que de cualquier argumento que se haya podido construir. El resultado es que muchas de las conclusiones a las que se llegan en estos estudios parecen surgir más de un acto de fe que del tratamiento formal de los argumentos. Incluso, cuando las descripciones son acertadas y de gran valía es difícil ir más allá con ellas". Y continúa: "es dentro de esta problemática donde tenemos que enmarcar el uso de SIG y de modelos matemáticos. Más concretamente, hacernos la pregunta de si es posible utilizar tales métodos para explorar, de una forma más rigurosa, conceptos y/o procesos identificados en estudios de paisaje de carácter interpretativo, nociones que suelen ser tratadas estricta y únicamente en narraciones".

Por tanto, a partir del modelo interpretativo de la localización espacial definido en los trabajos previos, se pueden factorizar diferentes criterios espaciales en un entorno SIG, mediante la creación de mapas raster que buscan aproximarse a las variables desde un punto de vista cuantitativo. Como indicamos anteriormente, la herramienta analítica principal que proponemos para investigar las propiedades y efectos de primer orden es el modelado predictivo, que permite, en último término, valorar en qué medida los yacimientos arqueológicos se encuentran distribuidos con relación a ciertas variables medioambientales.

Así, los criterios del modelo locacional se desmenuzan en factores que pudieron haber influido en las decisiones tomadas por las sociedades pasadas. M. García García (2015) define, en su tesis doctoral, un modelo de variables tripartito, estudiando la orografía, visibilidad y accesibilidad de ciertos asentamientos. En el caso del Megalitismo gallego, nosotros decidimos articular el análisis locacional en base a un modelo de factores dual sobre el que se estudiaron aquellas variables que, extraídas del modelo teórico previo, pudieron haber influido en el emplazamiento de estos monumentos, tal como vemos en la **Figura 36.**

En primer lugar, se hace hincapié en condicionantes físicos y orográficos, como la altitud, pendiente, prominencia topográfica, hidrología o geología. En segundo lugar, condicionantes que denominamos bajo el calificativo de potenciales, que engloban principalmente dos líneas de trabajo: el estudio del tránsito potencial y la relación de los monumentos con vías potenciales, y el estudio de la visibilidad. Por último, las variables de segundo orden, que se clasificaron, por motivos de exposición, bajo el calificativo de condicionantes culturales, con

ANÁLISIS LOCACIONAL MODELADO PREDICTIVO		Variables	Descripción
Factores de primer orden	*Condicionantes físicos y orográficos*	Altitud	Modelo Digital de Elevaciones de 25 m de resolución, construido a partir de datos LiDAR (sólo terreno).
		Pendiente	Inclinación máxima de la elevación en un punto determinado. Derivada del MDE.
		Hidrología	Red hidrográfica potencial, a partir de aquellos puntos cuyas características topográficas los hagan más proclives a haber funcionado como lugares de acumulación de agua.
		Geología	Cartografía en formato shapefile, digitalización de la Cartografía MAGNA del IGME.
		Zonas de humedal	Índice de humedad topográfica, calculado a partir de Boehner (*et al.* 2002).
	Condicionantes potenciales	Prominencia topográfica	Definida como la función de la altura diferencial entre un individuo y el entorno tal y como es percibida desde el punto de vista del individuo en cuestión (Llobera, 2001).
		Geomorfología	Caracterización topográfica en relación a las formas geomorfológicas del terreno.
		Tránsito potencial	Identificación de áreas que presentan mayor probabilidad estadística de ser transitadas, lo que las convierte en puntos clave o nodos teóricos en una red de tránsito peonil (Rodríguez Rellán, Fábregas Valcarce, 2015).
		Prominencia visual	Identificación de aquellas áreas del paisaje que son percibidas con mayor frecuencia.
Factores de segundo orden	*Condicionantes culturales*	Tradición	Valorar en qué medida existe una dependencia espacial entre los puntos (tendencia a la atracción o repulsión).

Figura 36. Criterios espaciales del modelo interpretativo del Megalitismo gallego definidos a partir de trabajo bibliográfico, junto a su factorización como variables raster susceptibles de ser analizadas en un Sistema de Información Geográfica (Carrero-Pazos 2017: 146).

las que podemos aproximarnos a las relaciones de dependencia entre los túmulos megalíticos, siguiendo lo que, por ejemplo, V. Villoch Vázquez (Villoch Vázquez 1995: 41-42) etiquetaba como tradición, “que llevaría a que un túmulo se sitúe donde ya existía otro anteriormente” (hipótesis ya presente en trabajos anteriores, como Vaquero Lastres 1991-1992: 173-174).

Finalmente, queremos puntualizar dos aspectos más teóricos que prácticos. Investigar fenómenos que podríamos definir como interpretativos (Gaffney, Van Leusen 1995), como la perceptibilidad de un monumento o la sonoridad de un paisaje, son aspectos complicados de cuantificar y cartografiar, aunque hoy se puedan modelizar más o menos objetivamente con un SIG (Wheatley 2000; Mlekuz 2004; Rennell 2012). No olvidemos, tampoco, que todo lo producido por un SIG es, en realidad, un modelo, y los resultados obtenidos deben ser

considerados probabilidades a investigar más profundamente en etapas posteriores (cfr. Wheatley 1995: 182-184).

Las variables que a continuación se analizarán con métodos SIG, pueden estudiarse individualmente mediante estadística univariada para comprobar su grado de correlación y, posteriormente, incluirse en un modelo de regresión logística multivariante. De la interacción de cada variable individual con las demás resultará o no como una buena variable predictora de los monumentos megalíticos y, por lo tanto, podremos deducir que influencia la distribución e intensidad de estos.

4.1.1. Factores de primer orden

a) Condicionantes físicos y orográficos

Quizás la variable más básica en cualquier estudio del paisaje sea la altitud. El análisis de la altitud se realiza en un SIG a través de un Modelo Digital de Elevaciones. Es necesario observar si existen tendencias altitudinales destacables en el emplazamiento de los sitios arqueológicos o si por el contrario cumplen el patrón altitudinal del área de estudio. Otra variable por inspeccionar es la pendiente, factor que se calcula en un SIG como la inclinación máxima de la elevación en un punto determinado, de ahí que en lenguaje de cálculo suela denominarse derivada de primer orden de la forma del terreno. La pendiente se mide calculando la tangente de la superficie y ésta, a su vez, se calcula dividiendo el cambio vertical en altitud entre la distancia horizontal. Por lo tanto, la pendiente es entendida de esta forma como un índice de cambio (inclinación) en el espacio, y puede variar según la dirección de la marcha (Conolly, Lake 2009: 252).[40]

El análisis de la pendiente es frecuente y harto conocido en los estudios arqueológicos (p. e. Warren, Ash 2000; Woodman 2000), ya que puede explicar o predecir la localización de elementos arqueológicos como la construcción de determinados sitios en lugares que rechazan pendientes pronunciadas. Esta variable es asimismo utilizada como factor para determinar la erosión (Burrough, McDonnell 1988), y relacionada con esto pueden determinarse áreas de relleno de sedimentos. Forma una parte relevante, también, en los estudios de tipo hidrológico, como los análisis de flujos.

En lo que respecta a la hidrología potencial, pese a que muchos estudios arqueológicos utilicen como base cartográfica la red hidrográfica actual y esto sea perfectamente válido, se suele descartar su uso dado que en los últimos 100 años la red hídrica de los territorios españoles se vio fuertemente alterada, con la desaparición de muchos cursos de agua y la modificación artificial de otros. Así, se recurre a la modelización de una red hidrográfica potencial a partir de aquellos puntos cuyas características topográficas los harían más proclives a haber funcionado como lugares de acumulación de agua.

Por otra parte, para medir la relación entre dichas zonas de acumulación de agua y los sitios arqueológicos se puede optar por un sencillo cálculo de distancias. La distancia calculada por un SIG suele ser euclídea, lo que significa que es calculada en línea recta. En nuestro trabajo doctoral optamos por un cálculo de coste de tránsito, entendido como el esfuerzo en tiempo

[40] Para el cálculo de esta variable existen varios algoritmos, y los más utilizados por los diferentes SIG son el método de Zevenbergen y el método de Horn (Burrough, McDonnell 1988: 192, para una explicación detallada).

necesario para recorrer la distancia topográfica entre un determinado curso de agua y un túmulo megalítico concreto. Esto nos permitió, por un lado, comprobar la existencia, o no, de relación entre los sitios arqueológicos y las fuentes acuíferas y, por otro, estudiar la facilidad o dificultad de acceso a los recursos hidrológicos en el paisaje.

El procedimiento que seguimos para calcular la red hidrográfica potencial se compone de varios estadios metodológicos. El primer paso necesario es eliminar puntos problemáticos del MDT, dado que el cálculo de acumulación de agua y direcciones es muy sensible a posibles incongruencias, como los denominados "fosos". Para ello, podemos utilizar la herramienta de eliminación de huecos en cualquier SIG al uso, como SAGA GIS (*Pit Remove*). La metodología que este software utiliza para eliminar los fosos, bien sean agujeros naturales en el paisaje o productos artificiales de la interpolación, es proporcionarles un valor de elevación congruente con sus celdas aledañas. Si no los corrigiésemos, incurriríamos en el error de considerarlos como sumideros, contenedores de agua que no son reales.

Una vez tenemos el MDE corregido, podemos crear una capa de drenaje de flujo (*flow direction*), que se concreta en calcular la dirección que seguiría una masa de agua artificial que discurriese por el Modelo Digital de Elevaciones. Los valores negativos indican que las celdas reflejan flujos procedentes de fuera de la región de estudio. La dirección que sigue este flujo permite observar la acumulación de varios flujos, es decir, analizar cuántas celdas del mapa drenan hacia otra *(flow accumulation)*.

Para el cálculo de ambos mapas, la dirección de flujo y la acumulación, podemos optar por dos tipos de métodos, el algoritmo de flujo *D8* y el *D-infinity* (Tarboton 1997).[41] El cálculo de flujo mediante el método *D-infinity* asigna un valor de dirección a cada celda basándose en la pendiente más pronunciada de una representación triangulada, consiguiendo resultados más ponderados y por lo tanto más ajustados a la realidad que con la aproximación de flujo D8.[42] Un paso analítico frecuente en estos momentos es convertir los valores del mapa de flujo acumulado a otros más manejables mediante una función logarítmica del valor absoluto de la acumulación, que puede realizarse sencillamente en *GRASS GIS* utilizando la herramienta *r.mapcalc*, mediante la siguiente fórmula: *log_accumulation = log(abs(accumulation)+1)*.

Desde aquí, quizás nos interese establecer cuál es el umbral a partir del cual el drenaje da lugar a un curso de agua. Este paso, según indica M. García García (2015: 751), debe realizarse comparando ciertas reclasificaciones del mapa de drenaje con la hidrografía actual, para establecer la máxima equivalencia posible entre la hidrología potencial y la actual. Por último, tendríamos la reclasificación de la hidrología resultante mediante el método de *Strahler*, en el que se asigna un orden numérico a cada corriente, cuyo valor se ve incrementado si dos cursos del mismo orden se juntan.

Por lo que respecta a la geología, para valorar su importancia como factor de primer orden suele utilizase la cartografía raster MAGNA del IGME (Instituto Geológico y Minero de España).

[41] No nos detendremos aquí en la explicación técnica sobre las diferencias principales entre ambos algoritmos. En todo caso, puede consultarse Burrough, McDonnell 1988: 194.

[42] El modelo que proponemos para el cálculo de una red hidrográfica potencial nos da una idea de la cantidad de agua que puede fluir por un área de estudio, aunque todo ello debe ser entendido como un modelo abstracto. Hay que tener en cuenta, además, que este tipo de cálculos no prestan atención a fenómenos de relevancia para la hidrología, como la evaporación, filtración, flujos subterráneos, etc.

Por otra parte, algunos autores han señalado una relación visual entre zonas de humedal (conocidas como brañas) y los megalitos (Méndez Fernández 1998; Santos Estévez 2008). Un buen trabajo de análisis locacional debería estudiar, entonces, si las zonas de humedal pudieron haber funcionado como un factor locacional, o si mantienen algún tipo de relación espacial con los sitios megalíticos. Modelizar esta variable supone hacerlo sobre un factor que no cuenta con cartografía básica que pueda orientar un análisis SIG, a la vez que este tipo de zonas suelen ser extremadamente cambiantes.[43] SAGA GIS permite crear un sencillo mapa de acumulación de flujo a partir de un índice de humedad topográfico (SAGA *Topographic Wetness Index*), que trabaja con la siguiente ecuación: *TWI= Ln ("f"/Tan ("s"))*, donde: *f* es la acumulación de flujo y *s* la pendiente. El mapa resultante es un raster de humedales, en el que las celdas representan una tendencia de escorrentía, considerando que un área con mayor humedad es más propensa a saturarse y formar un humedal. Los valores raster se categorizan mediante un índice de humedad constante relacionado con la topografía del terreno, y que puede ser utilizado para describir la distribución espacial de la humedad. Existen diversas formas de calcular este índice, aunque nosotros optamos por la variante propuesta por J. Boehner (*et al.* 2002) en SAGA GIS, que produce de forma rápida un mapa potencial de humedales (**Figura 37**).

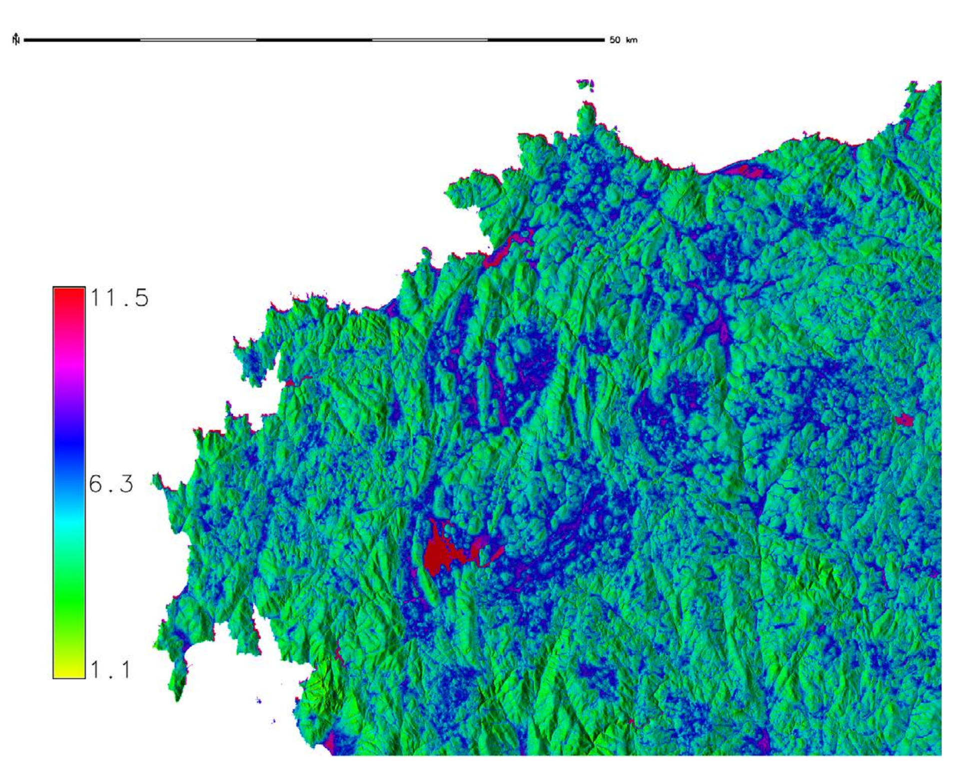


Figura 37. Cartografía de los humedales potenciales para el noroccidente de Galicia, a partir del cálculo de un índice de humedad topográfica. Nótese los valores más elevados (en rojo), que caracterizan el embalse de A Fervenza, sobre el curso medio del río Xallas.

[43] Algunos investigadores han intentado proponer otras soluciones, como la aportación de F. González Insua (2013), que elabora un mapa de zonas húmedas a partir de la toponimia presente en la cartografía 1:5000 y el catastro actual.

b) Condicionantes potenciales

En nuestro trabajo doctoral definimos toda una serie de condicionantes que denominamos potenciales, al no corresponderse con variables geográficas primarias, sino más bien reflexiones interpretativas procedentes de investigaciones anteriores. Un ejemplo es la localización de los megalitos en zonas de prominencia topográfica, factor propuesto principalmente por F. Criado Boado (1988). Para modelarla en un entorno SIG, optamos por seguir los trabajos de M. Llobera (2001, 2007), y su cálculo puede realizarse con *GRASS GIS* gracias a la extensión *r.prominence*.[44] Este módulo calcula la diferencia media entre una celda central y sus vecinas, siendo una aproximación a la robustez del terreno, pues en realidad examina las diferencias de la altitud dentro de una zona determinada (**Figura 38**).

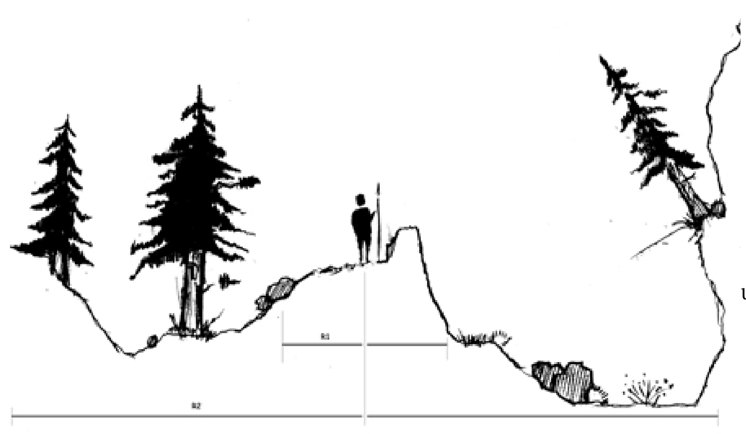


Figura 38. Esquema del cálculo de la prominencia topográfica, definida con relación a un radio específico definido por el usuario (Llobera 2001: 1007).

Una variante muy utilizada en los trabajos arqueológicos es el Índice de Posición Topográfica (TPI, en sus siglas en inglés), calculado a partir del módulo homónimo en el software *SAGA GIS*. Este índice se concreta, al igual que el caso anterior, en la comparación de la elevación de cada celda del DEM a la media de las elevaciones circundantes definidas por el usuario (Guisan *et al.* 1999; Weiss 2000; Wilson, Gallant 2000). El resultado es un mapa raster donde los valores positivos de prominencia del terreno son más altos que la media de sus inmediaciones (por lo tanto, hablamos de zonas elevadas). Los valores negativos representan localizaciones que son más bajas que sus inmediaciones, configurándose entonces como valles. Los valores que se encuentren cerca de cero se referirán a zonas planas (Weiss 2000) (**Figura 39**).

Para el estudio de la intensidad del tránsito potencial y su relación con los túmulos megalíticos, optamos por seguir los trabajos de P. Murrieta-Flores para el estudio del Megalitismo del sur peninsular (véase, por ejemplo, Murrieta Flores 2012), metodología planteada también, aunque con modificaciones, por M. Llobera (2015) para el estudio del conjunto megalítico de A Serra do Barbanza (Galicia).[45]

[44] https://grass.osgeo.org/grass64/manuals/addons/r.prominence.html

[45] Otros trabajos que también han contribuido con simulaciones de tránsito son los de White, Barber 2012; Verhagen 2013; Murrieta-Flores *et al.* 2014; Yubero-Gómez *et al.* 2015; Orengo, Livarda 2016.

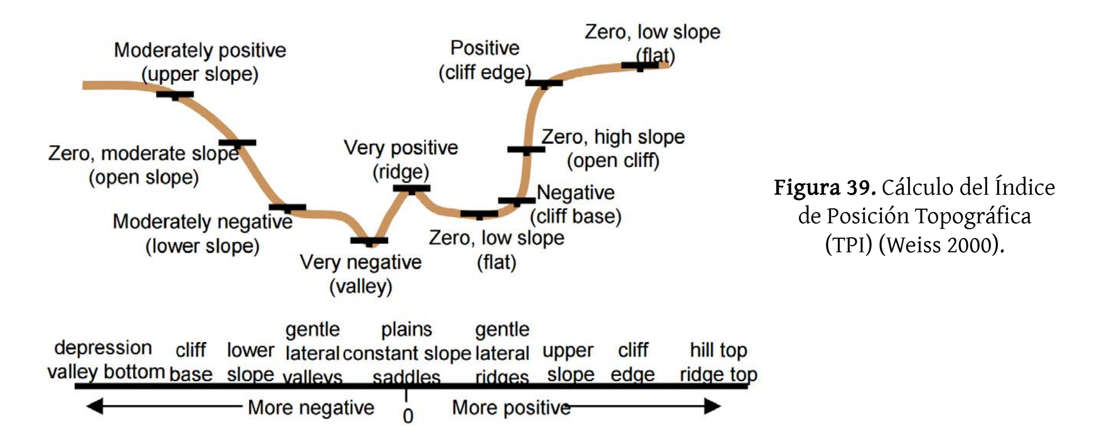


Figura 39. Cálculo del Índice de Posición Topográfica (TPI) (Weiss 2000).

Los modelos de movimiento humano desarrollados en un entorno SIG se concretan en modelos matemáticos aplicados a superficies raster, cuyas características son factores que determinan las condiciones del movimiento. De entre las técnicas más frecuentes, la ruta de menor coste (o *least cost path*) es una de las más utilizadas,[46] entendida como aquella ruta de tránsito generada entre dos puntos que corresponde al menor coste en tiempo o energía, dependiendo de determinados factores (Canosa-Betés 2016).

La metodología que explicaremos a continuación tiene por objetivo la creación de un modelo general de movilidad humana, partiendo de la generación de múltiples rutas de menor coste que comienzan y finalizan en los límites del área de trabajo. En cierta forma, lo realizado se concreta en una forma de movilidad desde un punto a la totalidad (*one to many*, en su terminología frecuente) aunque el cálculo se repite para un conjunto determinado de puntos (White, Barber 2012). En el trabajo de M. Llobera (2015: 84) se realiza un estudio del tránsito potencial a gran escala, y para ello podemos establecer los límites vectorizados del MDT como puntos de salida y llegada de las rutas. Este método nos permite obtener un conjunto de múltiples rutas de tránsito que, sumadas, proporcionan áreas con mayor probabilidad estadística de ser transitadas lo que, de facto, las convierte en puntos clave o nodos teóricos en una red de tránsito peonil (Rodríguez Rellán, Fábregas Valcarce 2015: 334) (**Figura 40A**).

A partir de aquí podemos calcular el esfuerzo en tiempo que tomaría llegar desde estas zonas de desplazamiento a los sitios arqueológicos estudiando la existencia o no de una relación (y su intensidad) entre la localización de los yacimientos y las rutas de tránsito potenciales (**Figura 40B**).

Para la creación del mapa de intensidad de tránsito potencial utilizamos el software *GRASS GIS*, comenzando con la creación de una superficie de fricción o coste de desplazamiento que expresa la permisividad al paso de cada celda del raster. Es frecuente que ésta incluya variables

[46] Siguiendo a J. Canosa-Betés (2016), la cuestión de la movilidad humada ha suscitado un interés particular desde los inicios de los SIG (Llobera 2000; Bevan 2011; Herzog 2013; Mlekuz 2014). Entiéndase la aproximación que aquí se plantea en un sentido exploratorio, y los resultados como una estimación estadística. Somos conscientes de las diferentes problemáticas que atañen al cálculo de las rutas de menor coste (véase Herzog 2013 como síntesis), habida cuenta de que es objeto de constantes modificaciones (Fairén 2004; Fábrega Álvarez, Parcero Oubiña 2007; Herzog, Posluschny 2011; Fovet, Zakšek 2014; Richards-Rissetto, Landau 2014) con un debate que llega a otras disciplinas, como por ejemplo la geográfica (por ejemplo, Yu *et al.* 2003).

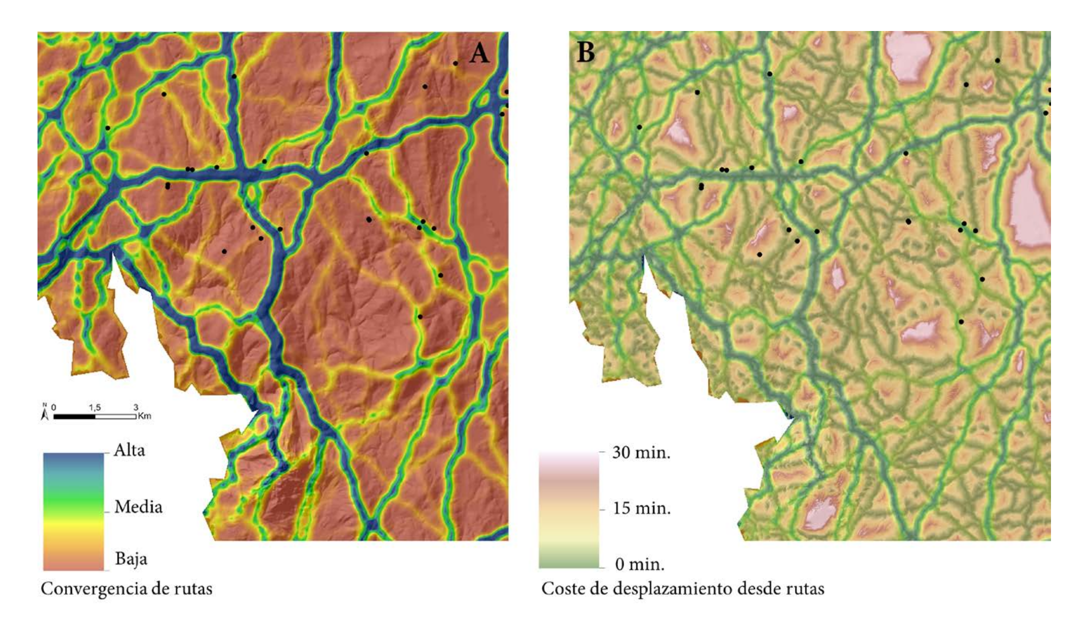


Figura 40. Densidad de rutas de menor coste para modelizar la intensidad de tránsito potencial, en la zona gallega de Costa da Morte (**A**) y coste de desplazamiento en tiempo desde las rutas de tránsito potencial al resto del paisaje (**B**).

como la topografía o hidrología, aunque también otras como la vegetación. Para mantener el modelo sencillo frecuentemente se opta por las dos primeras, creando un mapa de fricción con la herramienta *r.watershed*. El mapa raster generado debe ser reclasificado en cada caso concreto, para asignar valores de coste fuertes a zonas que deben bloquear el paso o al menos hacerlo difícil, como grandes caudales acuíferos, y otros más relativos a zonas por donde el tránsito debe fluir de forma sencilla, como zonas llanas de monte bajo.

El siguiente paso es crear una superficie acumulada del coste, en este caso con *r.walk*, partiendo desde un punto de inicio y llegando a uno de destino (denominado con frecuencia "algoritmo de difusión" o *spreading algorithm*).[47] Finalmente, utilizamos la herramienta *r.drain* para generar el vector con la ruta de menor coste.

Desde aquí, el proceso se repite para todo el conjunto de puntos y posteriormente se crea, con la herramienta *v.kernel* en *GRASS GIS*, un mapa raster de concentración de rutas en el que destacan zonas con densidad elevada. Por lo tanto, que potencialmente atraen el paso.

La relación entre estas zonas de movilidad potencial y los túmulos megalíticos puede calcularse a partir de la generación de un raster de coste de desplazamiento (en segundos) desde las propias rutas hacia el total del territorio, con la herramienta *r.walk*. El proceso analítico específico puede encontrarse en la **Tabla 2**.

[47] *R.walk* cuenta con ventajas frente a las implementaciones realizadas en otros software, porque, en primer lugar, utiliza la versión modificada de la "Regla Naismith" sobre el movimiento pedestre (Naismith 1892), coherente con los estudios fisiológicos de A. Minetti (*et al.* 2002), y en segundo lugar, consigue modelizar los costes anisotrópicos de forma adecuada, presentando los resultados en tiempo de desplazamiento (segundos) y no en unidades de coste no calibradas.

Cálculo de densidad de rutas óptimas en GRASS GIS			
Módulo utilizado	**Objetivo**	**Input**	**Output**
v.to.points	Descomponer el vector (línea) del área de estudio en puntos espaciados regularmente	Vector de línea del área de estudio	Vector de línea del área de estudio descompuesto en puntos regulares
r.watershed	Generar la superficie de acumulación de flujo	MDE	Raster de acumulación de agua potencial
r.walk	Generar la superficie de coste de movimiento	Origen: punto individual Destino: vector de puntos	Raster de coste de movimiento anisotrópico
r.drain	Generar los vectores de rutas óptimas	Raster de coste de movimiento anisotrópico (origen: vector de puntos)	Vector de rutas
r.drain se repite *n* veces			
v.kernel	Generar una superficie de densidad gaussiana kernel	Vector de rutas	Raster de densidad de rutas de tránsito óptimas

Tabla 2. Descripción metodológica del cálculo de densidad de rutas óptimas con *GRASS GIS* (a partir de Carrero-Pazos 2018a).

Por lo que respecta a la visibilidad, estamos ante uno de los factores que mayor atención ha recibido desde la investigación arqueológica con SIG y estadística espacial. Los estudios de visibilidad se desarrollaron a la par que los Sistemas de Información Geográfica evolucionaban, y adquirieron gran protagonismo historiográfico ya desde los inicios (Lake, Woodman 2003). Aunque existen muchos ejemplos que podríamos traer a colación, destacaremos el temprano trabajo de D. Fraser (1988) en el que estudiaba, de forma manual, la orientación de la visibilidad de los *cairns* de corredor de las Islas Orkney. Será en los 1990 y cambio de siglo cuando, no obstante, asistamos a una multiplicación de trabajos que se ocupan de desarrollar análisis de visibilidad con SIG, con las aportaciones, entre otros, de R. Bradley (*et al.* 1993), incluyendo análisis visuales manuales y un estudio de la visibilidad del terreno o, de mayor importancia en el contexto de este libro, el trabajo de D. Wheatley (1995) en el que introducía la noción de análisis de visibilidad acumulada (*cumulative viewshed analysis*) en Arqueología. Es también en este contexto en el que debemos destacar la aportación de M. Llobera (1996), que desarrollaba nuevos métodos para estudiar la prominencia de una localización en el paisaje. Años más tarde, este mismo autor trabajará sobre el concepto de paisaje visual, o *visualscape*, explorando modelos de ratio de cambio en la intervisibilidad entre sitios arqueológicos determinados (Llobera 2003, 2007). Otros trabajos reseñables pueden ser la aportación de M. Lake (*et al.* 1998), en la que se ocupa de estudiar los efectos de borde y analizar los requerimientos computacionales en análisis de visibilidad acumulada, o el trabajo de V. Gaffney (*et al.* 1996), una aproximación al estudio de la perceptibilidad, en una clara atenta por demostrar la aplicabilidad de los SIG a marcos de estudio postprocesuales.

Desde aquí, con el cambio de siglo asistimos al desarrollo de trabajos que comienzan a beneficiarse de los avances informáticos, como es el caso de las aportaciones que intentan modelizar la visibilidad considerando la vegetación (Gearey, Chapman 2006; Llobera 2007). Por su parte, M. Gillings (2009) nos proporciona un ejemplo del uso del análisis de visibilidad para modelizar la prominencia visual de un paisaje, o M. Lake y D. Ortega (2013), que desarrollan métodos que permiten el uso de la simulación de Monte Carlo para el análisis de cuencas visuales extensas.

Desde una perspectiva SIG, el análisis de visibilidad (o línea de visión, como a veces es referido) se concreta en determinar, desde un punto de vista cartográfico, si un par de puntos son inter-visibles o no (Conolly, Lake 2009: 296). La cuenca visual (*viewshed*, en inglés) de un determinado punto de visión son todas aquellas celdas raster sobre las que dicho punto es divisable. Podemos definirla como el conjunto de celdas raster desde las que un punto puede ser visto (*reciprocidad*). En *GRASS GIS*, podemos utilizar el comando *r. los*[48] para crear una cuenca visual de un punto de visión determinado.

Desde aquí, el resultado más inmediato de un análisis de visibilidad es un mapa binario, compuesto por ceros y unos, que nos indica aquellas zonas no visibles desde ningún punto de visión (0) y visibles desde uno o más puntos de visión (1). Esto puede hacerse, de forma sencilla, en *GRASS GIS*, mediante el módulo *r.viewshed*[49] (**Figura 41**).

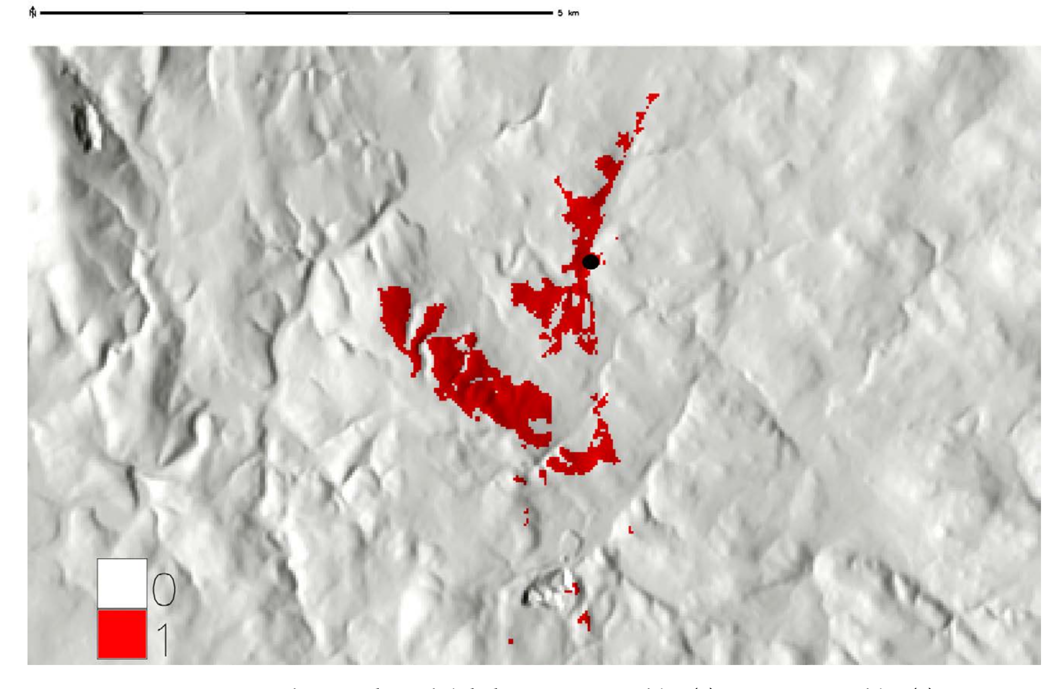


Figura 41. Mapa binario de visibilidad, con zonas visibles (1) y zonas no visibles (0), creado con el software *GRASS GIS*.

[48] http://www.ibiblio.org/pub/packages/gis/grass/grass62/manuals/html62_user/r.los.html
[49] https://grass.osgeo.org/grass82/manuals/r.viewshed.html

Aunque este tipo de análisis de visibilidad son conceptualmente bastante simples, la investigación ha desarrollado en los últimos años propuestas para afinar casuísticas concretas (Ogburn 2006; De Reu *et al.* 2011; Zamora Merchán 2011; Paliou 2013; Rášová 2014). Por ejemplo, si sumamos dos mapas de visibilidad tendremos un mapa de visibilidad múltiple, donde cada celda del mapa registra si ésta es visible o no desde al menos un punto de visión.

Finalmente, es necesario entender el concepto de visibilidad acumulativa (*cumulative viewshed*),[50] término introducido por D. Wheatley (1995: 173) para designar a la suma algebraica en un nuevo mapa de dos o más mapas de visibilidad binarios. Los valores de las celdas de un campo visual acumulado son números enteros entre cero y un número máximo teórico de puntos de visión (Conolly, Lake 2009: 297).[51]

Existen otros cálculos de visibilidad múltiple, más complejos, como la visibilidad en gradiente (*gradient viewshed*) o la visibilidad total (*total viewshed*) (Lake *et al.* 1998; Wheatley, Gillings 2000; Llobera 2003; Llobera *et al.* 2010). La visibilidad total se concreta en calcular la visibilidad desde todas las celdas del mapa raster, para posteriormente contar el número de celdas de todo el campo visual y luego registrar el número de celdas de otros mapas que son visibles desde allí (Conolly, Lake 2009: 298).[52] En *GRASS GIS* existe un comando para realizar este cálculo, *r.cva*[53] y su modificación actual, *r.viewshed.cva*.[54] Finalmente, se han desarrollado aproximaciones de interés hacia las zonas invisibles, es decir, aquellas que no son visibles desde determinadas localizaciones (Gillings 2015).

Existen también toda una serie de cuestiones concernientes a los análisis de visibilidad que también deben ser tenidas en cuenta, señaladas frecuentemente en la bibliografía al uso (véase Conolly, Lake 2009: 298 para una explicación detallada). Destacaremos ahora, por ejemplo, la importancia de especificar el software utilizado en las investigaciones. De hecho, tanto P. Fischer (1993) como P. Izraelevitz (2003) llaman la atención sobre este aspecto cuando indican que dependiendo del SIG que utilicemos, con diferentes algoritmos de cálculo, los resultados pueden ser distintos, aunque los datos sean los mismos.

Para el cálculo de la visibilidad topográfica absoluta podemos seguir las propuestas de M. Llobera (2006, 2007; Llobera *et al.* 2010) que permiten identificar, con mayor o menor detalle dependiendo del caso de estudio, aquellas áreas del paisaje que son percibidas con mayor frecuencia. El cálculo de la visibilidad total se define como "la visibilidad inherente a todas las localizaciones en un paisaje" (Llobera *et al.* 2010) y se realiza a través del cálculo de visibilidad de cada una de las celdas de un Modelo Digital del Terreno, lo que, dependiendo del caso, puede resultar en un elevado esfuerzo computacional.

El resultado de dicho cálculo se corresponde con zonas que son visualmente más prominentes, en el sentido de una mayor magnitud visual (Llobera 2003). Como este cálculo requiere procesos computacionales elevados, es frecuente realizar una aproximación a la visibilidad

[50] También denominado *times seen* por P. F. Fischer (*et al.* 1997).
[51] Sobre este respecto, Fels (1992) distingue dos tipos de análisis de visibilidad acumulativa, dependiendo de si se proporciona una altura determinada al sujeto de visión (*projective viewshed*) o al terreno (*reflective viewshed*).
[52] J. Lee y D. Stucky (1998) proponen dos variantes, según se incluya en el cálculo una altura para el observador (*viewgrid*) o para el terreno (*dominance viewgrid*).
[53] https://www.ucl.ac.uk/~tcrnmar/r.cva.html
[54] https://svn.osgeo.org/grass/grass-addons/grass7/raster/r.viewshed.cva/r.viewshed.cva.html

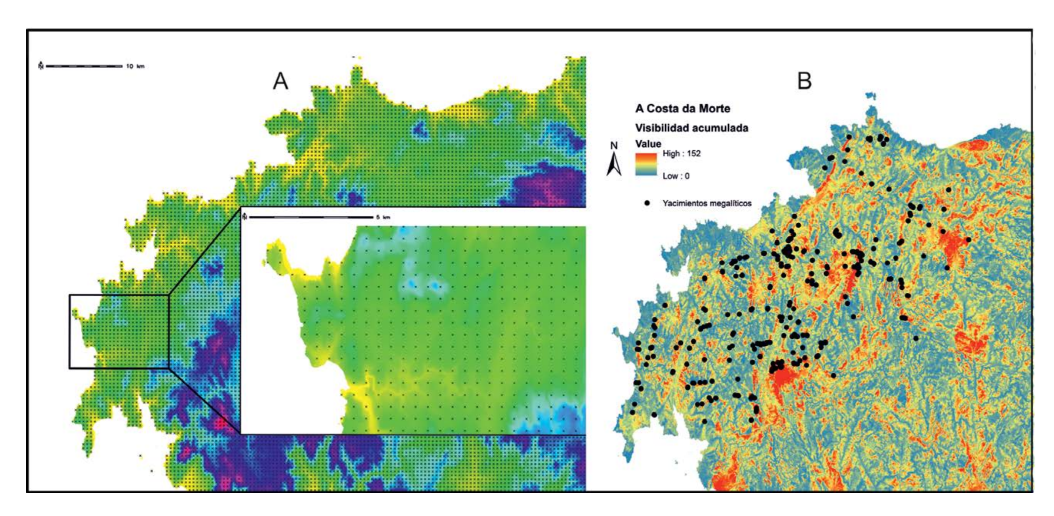


Figura 42. A: Cálculo de visibilidad acumulada para el occidente coruñés, basado en la creación de una malla regular de puntos situados a 250 m; **B:** Cálculo de visibilidad acumulada sobre 47.902 puntos para el área occidental coruñesa.

total, pero de forma muestral o selectiva. El método consiste en situar múltiples puntos distribuidos a lo largo del área de estudio y calcular la visibilidad acumulada desde cada uno de ellos. Para el establecimiento de los puntos podemos optar por crearlos de forma aleatoria (tal como hacen Llobera 2015 o Rodríguez Rellán, Fábregas Valcarce 2015), o generar una malla de puntos regulares por toda el área de estudio (espaciados a cierta distancia) y calcular, para todos ellos, la visibilidad acumulada. Así, a continuación, mostramos el resultado de una prueba para el caso del occidente de la provincia de A Coruña. Así, establecemos una malla regular de puntos regulares espaciados cada 250 metros (1 punto cada 250 metros) (**Figura 42A**). Ello supuso realizar el cálculo de la visibilidad acumulada sobre 47.902 puntos, consiguiendo un mapa de visibilidad acumulada del terreno bastante preciso (**Figura 42B**).

4.2. Modelado predictivo

El modelo predictivo es un método analítico que nos permite predecir el valor (la probabilidad de presencia) de una variable dependiente en un lugar no muestral utilizando una o más variables independientes. En Arqueología se han utilizado para conocer la probabilidad de presencia de yacimientos arqueológicos, sobre la base de una estimación cuantitativa de las características locacionales de sitios arqueológicos de un área prospectada (cfr. Conolly, Lake 2009: 238-239; Kvamme 1983, 1990; Judge, Sebastian 1988; Wescott, Brandon 1999).

El uso de los modelos predictivos en Arqueología ha sido acusado, al igual que los SIG, de determinista, por su excesiva dependencia en variables medioambientales a la hora de definir pautas de localización en sociedades del pasado (Gaffney, Van Leusen 1995). La dificultad de incluir variables de tipo cultural o temporal en los modelos predictivos ha sido, y continúa siendo en la actualidad, uno de los retos más importantes en la investigación con SIG, aunque las modernas técnicas de simulación estadística están permitiendo la renovación de los métodos hacia nuevos caminos analíticos (Verhagen 2018; Nüsslein *et al.* 2020). En este sentido, ciertos avances han sido realizados, por ejemplo en cuestiones

como la visibilidad o el estudio del movimiento o tránsito natural, con nuevas propuestas sobre cálculos de perceptibilidad, aunque las variables medioambientales sigan siendo las más utilizadas por su facilidad de medición y porque, en definitiva, algunas de ellas no han mudado drásticamente con el paso del tiempo, como el relieve, el drenaje y la permeabilidad del terreno, o la pendiente, entre otras (Kvamme 1983, 1990; Duncan, Beckman 2000; Warren, Asch 2000; Woodmann 2000).

Por otra parte, existen diferentes aspectos de los que, al utilizar el modelado predictivo, el investigador debe ser consciente. En primer lugar, es frecuente que yacimientos arqueológicos de diferentes períodos sean asumidos bajo un mismo arco temporal, al no contar con dataciones para la totalidad de una distribución analizada. Los análisis deben atender, entonces, a una incertidumbre temporal con el uso de modelos probabilísticos (Crema 2015). Por otra parte, tampoco se suele explorar con suficiente atención las relaciones univariantes existentes entre determinados sitios arqueológicos y las variables físico-geográficas, tanto positivas como negativas. Este ejercicio es muy recomendable antes de proceder a su combinación mediante un análisis multivariante, para evitar problemas de multicolinealidad, es decir, que dos variables muestren una fuerte relación entre sí. Por otra parte, es frecuente asumir estacionariedad en nuestra área de estudio. Es decir, a la hora de analizar una distribución de yacimientos determinada, normalmente se asume que el proceso espacial que está detrás de la confección de la distribución de propios puntos ocurre a lo largo y ancho del espacio analizado, a pesar de que esto es extremadamente inusual en la mayor parte de los fenómenos arqueológicos.

Los modelos predictivos pueden dividirse en tres tipos de categorías. En primer lugar, aproximaciones de densidad de puntos, aquellas que no realizan presunciones sobre la preferencia de localización de sitios. En segundo lugar, aproximaciones inductivas, aquellas que se basan en puntos conocidos con características locacionales que pueden ser generalizables al conjunto de la población y, en función de estos criterios, los modelos son determinísticos a nivel ambiental. Por último, tenemos aproximaciones deductivas, aquellas que pretenden responder a la pregunta de por qué los puntos están localizados en lugares específicos (Nakoinz, Knitter 2016).

El procedimiento de construcción de un modelo predictivo se define con relación a cuatro etapas, como son la recopilación de datos, el análisis estadístico, la aplicación del modelo y su validación (**Figura 43**).

Siguiendo a J. Conolly y M. Lake (2009: 240), se parte del presupuesto de que es posible diferenciar entre áreas del paisaje con evidencia de ocupación (es decir, “yacimientos”), y áreas del paisaje que no poseen dicha evidencia (puntos que identifican zonas sin yacimientos), en función de uno o más atributos del paisaje. Estas áreas se pueden establecer, por ejemplo, mediante muestreos aleatorios o de agrupaciones del paisaje. Este es un punto clave, porque en el ejercicio analítico podemos considerar un área que no posee yacimientos como una que sí los posee, o determinar zonas que no poseen sitios arqueológicos pero que en realidad hayan sido mal prospectadas y sí los contengan.

Por otro lado, la muestra de datos se suele dividir en “muestra de entrenamiento”, para crear el modelo, y “muestra de control”, que se utilizará para validarlo. Una práctica común es dividir

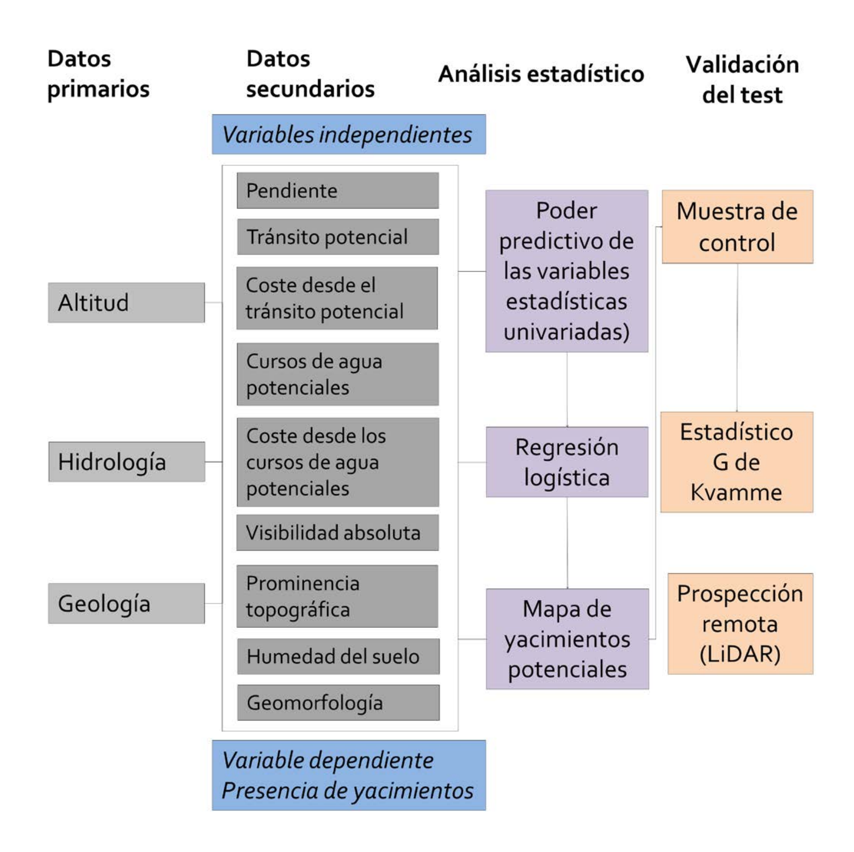


Figura 43. Metodología del modelado predictivo utilizado por M. Carrero-Pazos (2017: 181), a partir de las propuestas de J. Conolly y M. Lake (2009).

la base de datos inicial en un 50-40% y reservar, así, una muestra de control representativa del total de la población.

El primer estadio analítico debe estar encaminado a caracterizar el potencial predictivo de las variables elegidas, para lo que se suele utilizar estadística univariada. Antes de seleccionar los atributos, es interesante investigar si alguno de ellos puede sustituir a otro en la predicción de la presencia o ausencia de yacimientos, lo que se suele denominar "variables de confusión" (*confounding variables*), cuestión que se soluciona con una prueba de multicolinealidad (Woodman 2000: 452). Por otra parte, también debemos prestar atención a la interacción entre los atributos. Si una variable modifica la relación entre otras dos variables y la ausencia o presencia de yacimientos, aquella debe ser eliminada.[55]

[55] Un ejemplo claro de esto ya fue expuesto por P. E. Woodman (2000), quién estudiando el asentamiento mesolítico de Islay (Hébridas interiores, Escocia), descubrió que existían pocas posibilidades de localizar yacimientos en lugares donde escaseaba el agua y situados a gran altitud, con una mayor probabilidad de hallarlos en lugares áridos pero situados a baja altitud. La solución en este caso fue sustituir estas variables por una híbrida que intente representar la naturaleza de dicha interacción.

Posteriormente es necesario estudiar la relación de la distribución de datos observados con cada una de las variables medioambientales, utilizando una prueba univariante como la regresión lineal. Del análisis, el modelo descartará alguna variable y propondrá como relevantes otras tantas.

La técnica estadística más utilizada para crear el modelo predictivo es el análisis multivariante de regresión logística (Stopher, Meyburg 1979; Aldrich, Nelson 1984; Menard 2001). Se ocupa de estudiar la relación entre una variable dependiente y una o más variables independientes, denominadas covariables (regresión múltiple o generalizada). En concreto, trata de medir cómo los datos se ajustan a una relación modelada buscando, con ello, acoplar los datos a una curva de probabilidad en forma de *S*, lo que permite que las probabilidades predictivas de presencia de yacimientos cambien rápidamente entre probabilidad baja y alta (**Figura 44**).

La relación entre dos o más variables cuantitativas puede ser expresada en forma de grados de correlación (correlación positiva -si una variable aumenta, la otra también-, correlación cero -no existe relación entre las variables-, y correlación negativa -una variable aumenta cuando la otra disminuye-), o también es frecuente en términos de fuerza de la relación. Para ello, puede utilizarse el coeficiente de correlación *r* de Pearson. Este estadístico devuelve un valor que varía desde 1 (correlación positiva perfecta), hasta -1, 0 (correlación negativa perfecta)[56] (**Figura 45**).

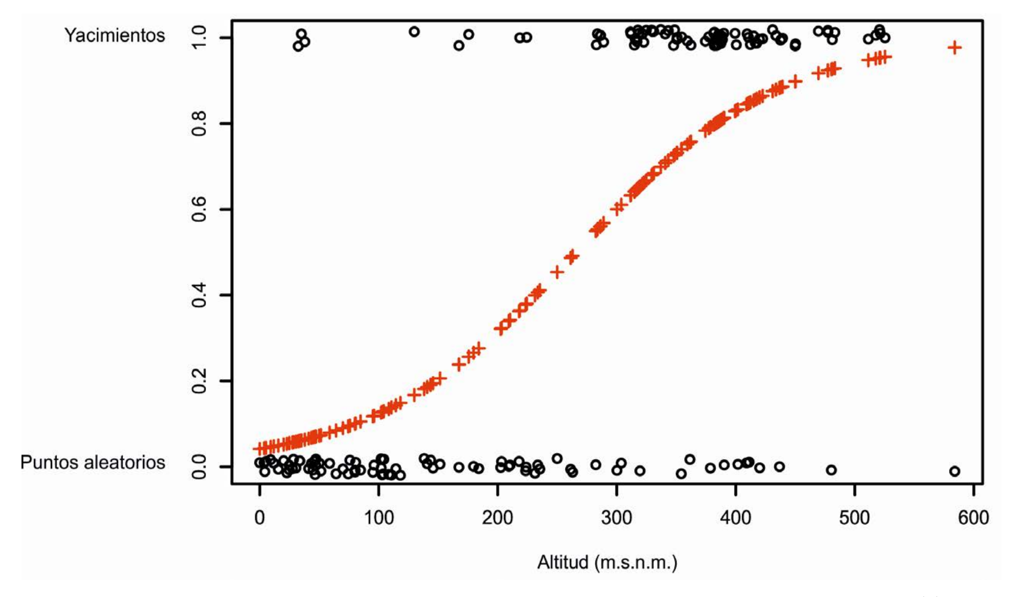


Figura 44. Ejemplo de regresión logística lineal, con probabilidad de sitios arqueológicos (1) y zonas identificadas como áreas sin yacimientos (0) con respecto a determinadas altitudes.

[56] El coeficiente se visualiza a menudo en forma de recta de regresión con un r^2, conocido como coeficiente de determinación. En su caso, oscila entre 0 (no existe correlación) y 1 (correlación perfecta). Lo interesante del coeficiente de determinación es su indicación predictiva de la variable independiente, y puede determinarse como la proporción de la variación de los valores de la variable *y* que están determinados por la variable *x*. Así un coeficiente de correlación de 0,96 se convierte en un coeficiente de determinación r^2 de 0,88, lo que indica que el 88% de la

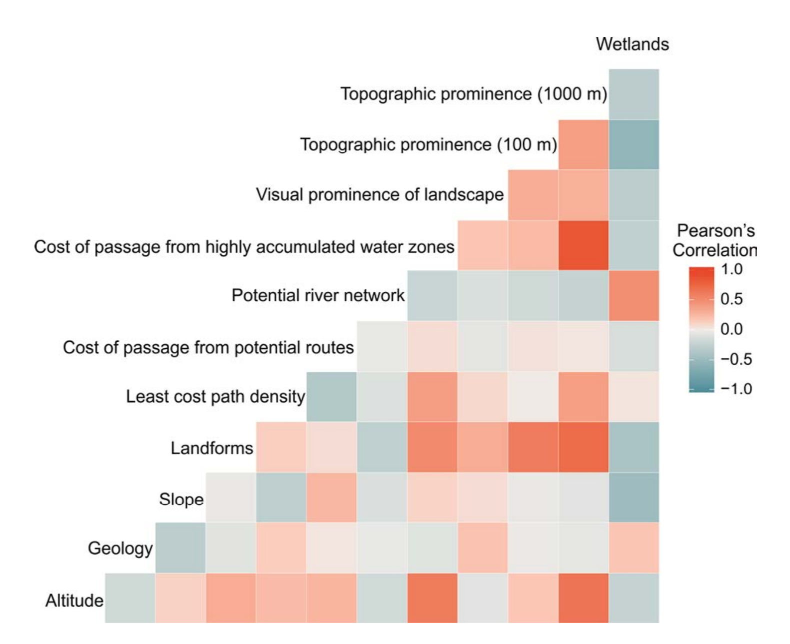


Figura 45. Ejemplo gráfico de correlación de Pearson, con diversas variables raster (Carrero-Pazos 2018b: 545).

El modelo de regresión logística generalizado (más conocido por sus siglas GLM, *Generalised Linear Model*), puede realizarse en *R Statistics* gracias a la función *glm*.[57] La ecuación de regresión es expresada como: $V=\alpha+b_1 x_1+b_2 x_2+\cdots+b_n x_n$ donde α es una intersección con una serie de coeficientes $b_1, b_2,..., b_n$ que determinan la ponderación aplicada a cada uno de los *n* atributos $x_1, x_2,..., x_n$. El primer paso es, por tanto, calcular el valor de *V* para cada celda raster (*i*), y, en un segundo momento, convertir este valor en probabilidad de presencia de yacimientos por cada celda (p_i), aplicando la siguiente ecuación (Haining 2003: 262): $p_i=V_i/(1+exp(V_i))$

Por otra parte, podemos, además, introducir un criterio formal de contrastación para obtener aquella combinación de variables que prediga, con mayor exactitud, la distribución de sitios arqueológicos. Se trata, en términos simples, de depurar el modelo de variables inicial para escoger aquellas que consigan predecir con mayor detalle la distribución de sitios arqueológicos. Uno de los métodos que puede utilizarse se denomina criterio de información Akaike, y es una medida de la calidad relativa de los modelos estadísticos (Baddeley *et al.* 2015: 335-336). En este caso, el criterio Akaike ponderará diferentes modelos, quitando y añadiendo las variables del modelo inicial hasta conseguir aquella combinación más exacta. En términos generales, cuanto menor sea el valor AIC, más preciso será el modelo (Bevan *et al.* 2013). Puede calcularse en *R Statistics* gracias a la función *stepAIC()* del paquete *MASS* (Venables, Ripley 2002).

El resultado de un modelo de regresión generalizado es similar a la **Tabla 3**, en la que pueden apreciarse, a la izquierda, las variables de estudio, junto a un valor *p* en la columna Pr(>|z|).

variación de *y* puede explicarse por las características de *x* (Conolly, Lake 2009: 207). Este modelo es lineal y bivariado, pues hace referencia a dos variables, y nos permite realizar predicciones de *y* para cualquier valor de *x*, dada la fórmula *y=α+bx*.

[57] https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/glm

Coeficients	**Estimate**	**Std.Error**	**Z value**	**Pr(>|z|)**	
(Intercept)	-3.4201714	0.6468606	-5.287	1.24e-07	***
Altitud	0.0137736	0.0023533	5.853	4.83e-09	***
Densidad de tránsito potencial	0.0657404	0.0172721	3.806	0.000141	***
Hidrografía potencial	-0.0450663	0.0214592	-2.100	0.035721	*
Coste de tránsito desde zonas de acumulación de agua	-0.0017089	0.0009755	-1.752	0.079794	.
Signif. codes: 0 ‘*’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1**					

Tabla 3. Ejemplo del resultado de un modelo de regresión logística multivariante. La columna Pr(>|z|) proporciona un valor *p* indicando aquellas covariables que son significativas.

stepAIC Function			
	Df	Deviance	AIC
<none>		107.79	117.79
Hidrografía potencial	1	114.56	122.56
Densidad de tránsito potencial	1	137.62	145.62
Altitud	1	167.59	175.59

Tabla 4. Resultado de la aplicación del criterio de información Akaike, que permite obtener la mejor combinación de variables iniciales posible.

En la **Tabla 4** podemos ver el resultado de la aplicación del criterio Akaike, que nos indica que la mejor combinación de variables es aquella compuesta solamente por tres. En el caso propuesto, por ejemplo, se ha eliminado el coste de tránsito desde zonas de acumulación de agua potencial. Nótese también que, en el modelo de regresión generalizado, esta variable poseía valores *p* no significativos.

Con el modelo de regresión generalizado ya depurado podemos calcular el modelo predictivo. Este paso se concreta en la creación de un mapa raster con los valores predictivos de cada covariable, sumando el valor estimado del *intercept* a la multiplicación de las variables y su valor estimado, consiguiendo, con ello, una superficie de probabilidad relativa. El valor predictivo de cada celda se computará utilizando la ecuación, $logit(p)=\alpha+\beta_{1_{x1}}+\beta_{2_{x2}}+\dots+\beta_{n_{xn}}$, donde α es el *intercept* (el valor predicho de la variable dependiente cuando todas las variables independientes son 0), β los coeficientes y χ las covariables. Esto nos proporciona un algoritmo de la probabilidad de presencia de yacimientos dividido por la probabilidad de ausencia de yacimientos. Posteriormente, para calcular la probabilidad relativa de la existencia de un yacimiento en una localización particular, se aplica la siguiente ecuación:

$$p_i = \frac{V_i}{1+\exp(V_i)}$$

El resultado final será un mapa raster que representa, con una probabilidad de entre 0-1,[58] aquellas áreas que, con un determinado valor de significación estadística, pueden contener sitios arqueológicos (**Figura 46**).

Finalmente, un modelo predictivo debe ser validado. La forma más utilizada se concreta en hacer uso de la muestra de control que se había reservado durante el proceso de creación del modelo, tal y como se observa en la **Figura 46**. De esta manera, se calcula cuántos sitios arqueológicos de la muestra de control caen dentro del área donde existe un valor predictivo alto. Si, por ejemplo, 72 yacimientos de un total de 90 estudiados caen dentro del área con más de un 70% de valor predictivo, diremos entonces que el modelo predice correctamente la presencia en un 80% de las veces. Sin embargo, existen toda una serie de consideraciones sobre las que conviene detenerse, como que la predicción es probabilística, y generalmente sólo tiene sentido hablar correctamente de un modelo predictivo con un cierto nivel de probabilidad, entre 0 y 1. Además, los modelos tienden a ser más precisos con probabilidades bajas, y menos exactos con probabilidades altas (Conolly, Lake 2009: 244).

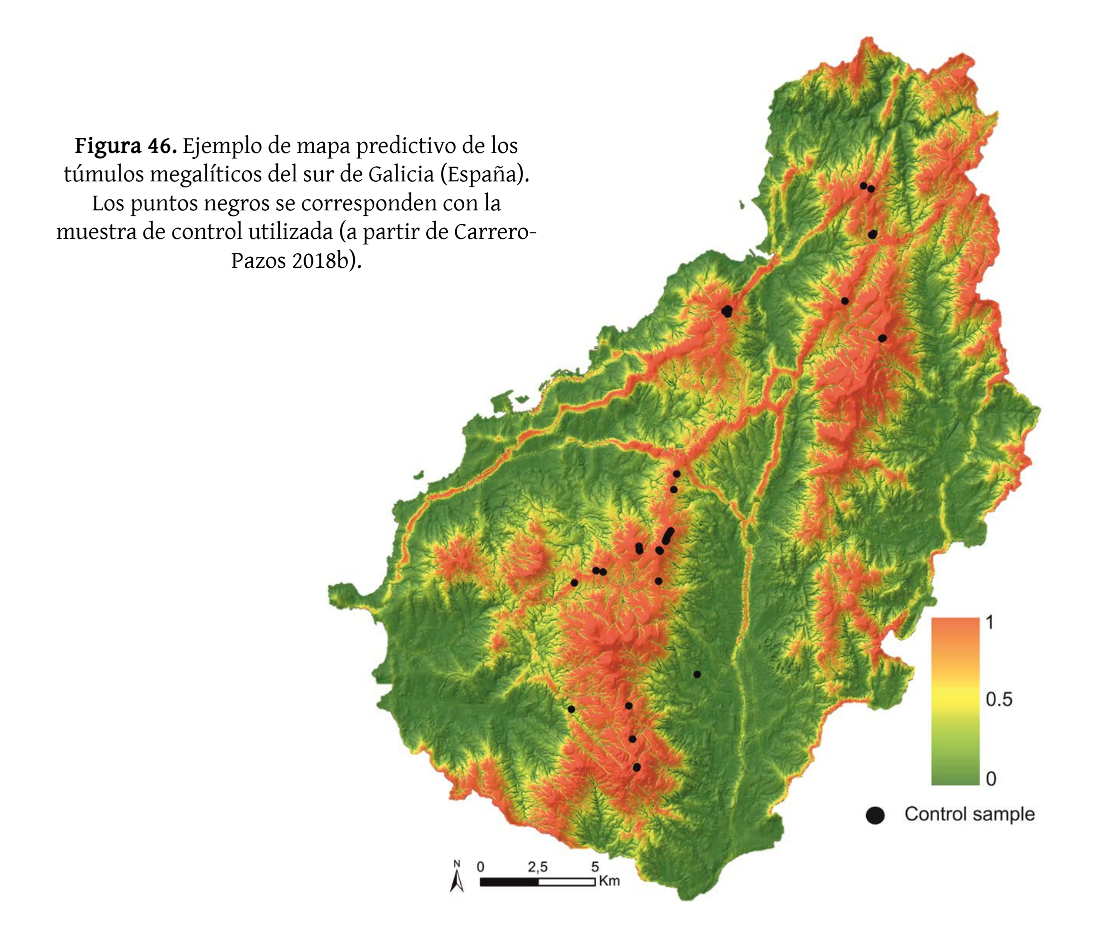


Figura 46. Ejemplo de mapa predictivo de los túmulos megalíticos del sur de Galicia (España). Los puntos negros se corresponden con la muestra de control utilizada (a partir de Carrero-Pazos 2018b).

[58] Mapa potencialidad arqueológica = (exp(Superficie_predictiva))/(1+(exp(Superficie_predictiva)))

Por otro lado, la distribución de puntos creada que no se corresponde con sitios arqueológicos, es clave. A menudo es posible especificar una probabilidad de aparición de yacimientos tan baja que toda la muestra observada cae de hecho dentro del área donde la predicción indica su presencia, pero casi siempre habrá un gran número de "no yacimientos" que también entren en esa misma área, con lo que el modelo es inexacto en la predicción de la ausencia de yacimientos arqueológicos.

Existen otros métodos que se pueden utilizar para continuar estudiando la exactitud de un modelo predictivo. Una primera opción es la producción de curvas de predicción correcta en porcentajes acumulativos para yacimientos y "no yacimientos", aunque el método más usado fue desarrollado por K. Kvamme (1988: 329) a finales de los años 1980, denominado *Kvamme's gain statistic* (valor de "ganancia" *G* de *Kvamme*). En términos estadísticos, se expresa por, $G=1-(\frac{S}{O})$ donde *S* es el porcentaje del total del área en la que los yacimientos están predichos; *O*, el porcentaje de los sitios observados que están dentro del área predictiva, lo que nos proporciona un valor de ganancia, o *G*, que se aproximará a 1 (alta utilidad predictiva), 0 (sin utilidad predictiva), o -1 (el modelo predice lo contrario de lo que supone).

La propiedad más importante de esta medida estadística es que distingue un modelo correcto, pero de relativa utilidad, de un modelo claramente menos correcto, pero más útil.[59] Así, como bien escribían J. Conolly y M. Lake (2009: 245), es probable que un modelo que predice correctamente un 80% de los yacimientos y predice la presencia de yacimientos en más de un 70% del paisaje no sea muy útil, lo que se refleja en una baja ganancia.

La mayor parte de los autores consideran que un modelo con *G* por encima de 0,5 debe ser aceptado (Ebert, Kohler 1988; Ebert 2000; Ducke, Munch 2005; Graves 2011: 641).

En definitiva, la composición y precisión de un modelo predictivo dependerá de lo rigurosos que seamos a la hora de escoger, discutir y contrastar las variables de partida, aspectos que se deben tener en cuenta a la hora de definir los objetivos de la investigación. Uno de ellos puede ser, por ejemplo, utilizar un modelo predictivo para encontrar nuevos yacimientos arqueológicos; otro, establecer aquellas variables que predicen, de forma correcta, una distribución de sitios arqueológicos, por lo que, en último término, permiten analizar el efecto que dichas variables de primer orden ejercen sobre la distribución arqueológica.

Finalmente, para estudiar el efecto que las variables de primer orden ejercen sobre el patrón espacial mostrado por un conjunto de sitios arqueológicos, podemos utilizar una función de correlación par, en la que los puntos que componen el conjunto de simulaciones aleatorias consideren, al distribuirse, el modelo de regresión generalizada creado anteriormente (**Figura 47**).

[59] Otra opción que está siendo explorada en la actualidad es el estadístico *J* de Youden (véase al respecto https://matthewdharris.com/2018/04/16/part-1-youdens-j-as-a-replacement-for-the-kvamme-gain-metric-in-the-evaluation-of-archaeological-predictive-models/).

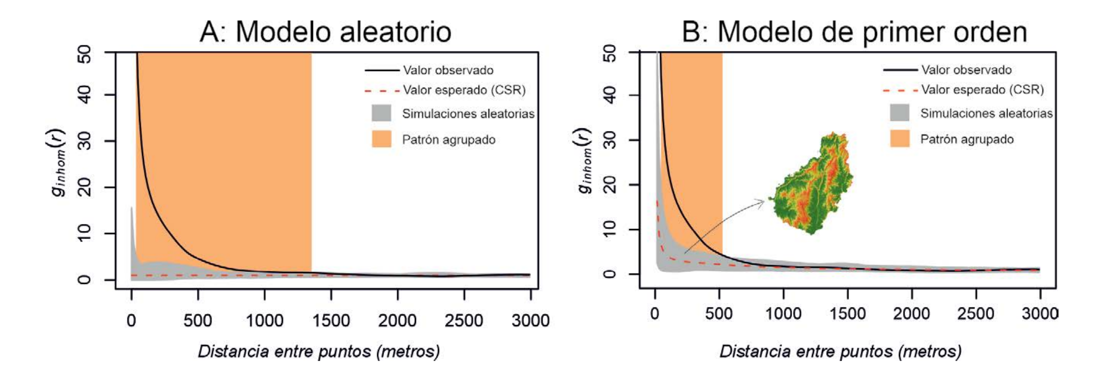


Figura 47. Modelos de relaciones espaciales entre túmulos megalíticos (línea negra) y simulaciones aleatorias (bandas de gris claro). El sombreado naranja ilustra tendencias de agrupación significativas. **A:** Función de correlación par en la que las simulaciones aleatorias están creadas bajo condiciones *Poisson* de aleatoriedad espacial completa (CSR). **B:** Función de correlación par en la que las simulaciones aleatorias están creadas teniendo en cuenta los efectos de primer orden (modelo de regresión logística).

En la **Figura 47A**, los puntos que componen el conjunto de simulaciones aleatorias están creados bajo parámetros de aleatoriedad espacial completa (CSR), lo que significa que estamos comparando la distribución de sitios arqueológicos frente a un modelo puramente aleatorio. De su análisis visual se puede concluir la existencia de agrupación en los datos hasta los 1.400 metros, aproximadamente. Es decir, los sitios arqueológicos en cuestión se distribuyen de forma agrupada, formando clústeres por el paisaje, y esto no es el resultado del azar.

En segundo lugar, podemos analizar las tendencias introduciendo los efectos de primer orden en la función de correlación par. En la **Figura 47B**, las simulaciones aleatorias de la función han sido generadas teniendo en cuenta el modelo de regresión logística generalizado, es decir, en realidad estamos condicionando las simulaciones aleatorias para su ajuste al modelo de regresión. De esta forma, controlamos los efectos de primer orden en el análisis de los patrones espaciales y, pese a introducir los efectos de primer orden en las simulaciones, se observa cómo continúa existiendo agrupación a escala local (0-500 metros) residual. Es decir, el patrón agrupado que muestran los sitios arqueológicos (en este caso, túmulos megalíticos), no puede ser explicado ni por la aleatoriedad ni por el efecto de las variables de primer orden.

Estos resultados pueden interpretarse en dos sentidos. En primer lugar, puede ser que existan variables de primer orden que no hayan sido incluidas en el estudio pero que sean relevantes en la localización de los sitios arqueológicos en cuestión o, en segundo lugar, que sean los propios sitios arqueológicos y sus relaciones de dependencia espacial las que expliquen este patrón agrupado a escalas locales (factores de segundo orden).

En definitiva, una vez hayamos definido las variables de estudio, procedido a su rasterización con un trabajo SIG y contrastado su importancia con respecto a la distribución de sitios arqueológicos utilizando, por ejemplo, un modelo de regresión logística generalizado, habremos modelizado, satisfactoriamente, la influencia de las variables de primer orden sobre la distribución de sitios arqueológicos. Si el modelo de variables de primer orden consigue explicar en su totalidad el patrón de puntos observado, habremos de parar aquí,

argumentando entonces, con criterios interpretativos, por qué esas variables específicas son relevantes y consiguen explicar el patrón de distribución de los sitios arqueológicos analizados. De no ser así, como ocurre en el caso anterior, es decir, de existir todavía ciertos patrones observados que no pueden ser explicados solamente por la acción de las variables de primer orden estaremos entonces en condiciones de analizar los factores de segundo orden. Tema del que nos ocuparemos en el siguiente capítulo.

Capítulo 5.

Simulación estadística para el estudio de patrones culturales

5.1. Factores de segundo orden. La interacción entre los puntos

Los factores de segundo orden se ocupan del análisis de la dependencia espacial entre los puntos. Ya hemos explicado que pueden estudiarse mediante las funciones K y sus derivadas, aunque, siguiendo el capítulo anterior, utilizaremos la función de correlación par para continuar investigando los patrones de localización del fenómeno tumular y megalítico gallego (véase Bevan *et al.* 2013; Spencer, Bevan 2018; Carrero-Pazos *et al.* 2019, Bevan 2020, para ampliar las explicaciones que siguen).

Recordemos que lo primero que debemos realizar es definir y explorar, de forma precisa y rigurosa, la influencia de los efectos de primer orden en el patrón de puntos. Es bien conocido que el fenómeno megalítico mantiene patrones agrupados en el paisaje, conformando cementerios que suelen estar en uso durante cientos de años, incluso milenios. Se espera entonces que los monumentos muestren un patrón agrupado, algo que confirmamos en nuestro conocimiento arqueológico previo y de forma cuantitativa a través de una función de correlación par (**Figura 48**), que muestra una tendencia agrupada a escala local que no puede ser explicada ni por la aleatoriedad ni por los efectos de primer orden. Cabe pensar, por lo tanto, que sea la dependencia espacial de los propios puntos, los megalitos, la que está provocando este patrón agrupado a escalas locales, es decir, que la localización de un túmulo megalítico haya atraído el emplazamiento de otros.

Para analizar los efectos de segundo orden podemos ajustar las simulaciones aleatorias a modelos estadísticos ya conocidos, hecho que nos permitirá considerar los efectos de segundo orden en su cálculo, es decir, simular relaciones de dependencia espacial entre los sitios arqueológicos (por ejemplo, fuerte agrupación entre ellos). En este sentido, existen muchos modelos de patrones estadísticos que pueden aplicarse dependiendo del diseño de la investigación (Nakoinz, Knitter 2016). Uno de los más utilizados es el modelo de interacción de área (Baddeley, van Lieshout 1995), que genera patrones de inhibición y agrupación con referencia a un *buffer* creado para cada punto de la distribución, y ha sido muy utilizado en Arqueología porque permite simular una noción de territorialidad (Bevan *et al.* 2013).

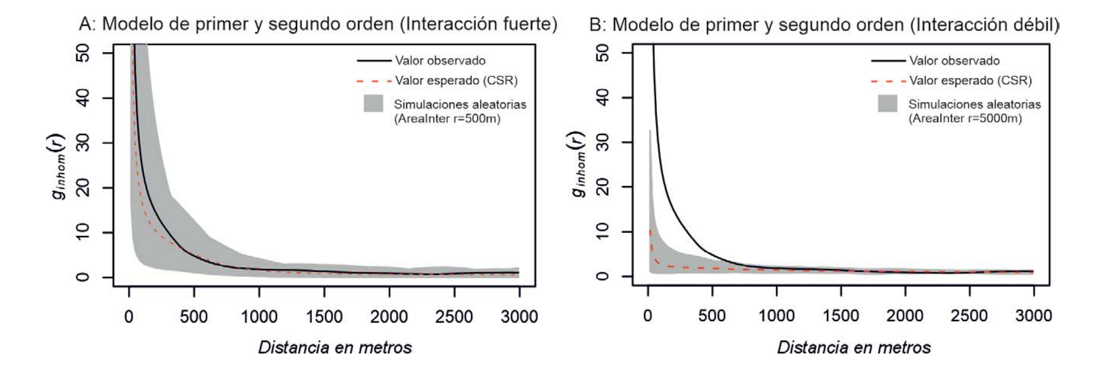


Figura 48. Modelos de relaciones espaciales de primer y segundo orden entre túmulos megalíticos (línea negra) y simulaciones aleatorias (bandas grises). En ambos casos las simulaciones aleatorias fueron creadas teniendo en cuenta los efectos de primer orden (regresión generalizada), junto a una interacción fuerte entre los puntos (r = 500 m) (**A**) y débil (r = 5000 m) (**B**).

C. Rodríguez Rellán y R. Fábregas Valcarce (2015) denominan acertadamente a este proceso analítico como "mapa de atracción", ya que en realidad lo que se genera es un modelo en el que cada punto posee un radio de atracción, como si de una noción de territorialidad se tratase. De esta forma, podemos crear un conjunto de simulaciones aleatorias que tengan en cuenta los efectos de primer orden pero que, a su vez, fuercen a los puntos que las componen a mantener valores de interacción fuertes o débiles entre ellos, y comparar los resultados con la tendencia de los sitios arqueológicos.

Los resultados del cálculo, presentados en la **Figura 48**, permiten observar que ahora la distribución de túmulos (línea negra) cae bajo el intervalo de confianza influenciado por los efectos de primer orden y una interacción fuerte entre los puntos, por lo que el modelo de interacción se ajusta a los datos arqueológicos (**Figura 48A**). Por otra parte, si utilizamos un radio de interacción más amplio, como 5000 metros (**Figura 48B**) -por lo tanto, creando relaciones de atracción débiles entre los puntos-, observamos que la tendencia de los datos arqueológicos no se ajusta al intervalo de confianza (banda aleatoria), con lo que, de esta forma, descartamos que la agregación de túmulos observada a escalas locales sea producto de una interacción débil entre ellos.

Recapitulando, mediante este ejercicio de simulaciones aleatorias y su comparación con la tendencia de la distribución observada (los sitios arqueológicos), hemos podido modelar los efectos que las variables medioambientales tienen en los patrones espaciales de un grupo específico de túmulos megalíticos gallegos, así como las relaciones de dependencia espacial entre ellos, en este caso, de fuerte interacción. En términos arqueológicos estos monumentos mantienen patrones de agrupación tanto a escalas generales como locales. Esto es algo harto conocido, aunque hemos conseguido definir cuál es la escala de dicha agrupación, y qué forma tiene: la tendencia espacial de agrupación observada a escala general puede entenderse, a juzgar por los análisis mostrados en este capítulo y el anterior, a partir de ciertas variables medioambientales, como pueden ser la proximidad a áreas de tránsito por el terreno o la localización sobre ciertos rangos altitudinales. Estas variables no parecen jugar un papel

relevante en escalas locales, en las que la razón principal de la localización parece ser una interacción fuerte entre los túmulos, es decir, que un túmulo se localizaría siempre en las proximidades de otros preexistentes, en conjunción con una tradición funeraria arraigada durante milenios.

TERCERA PARTE:

CIENCIA ABIERTA EN ARQUEOLOGÍA. CASOS DE ESTUDIO

Capítulo 6.

Ciencia arqueológica en acceso abierto y transparente

La disciplina arqueológica se enfrenta, en la actualidad, a nuevos retos epistemológicos, en parte forzados por el aluvión de datos, técnicas y métodos analíticos de los que disponemos, quizás ya más digitales que analógicos (Bevan 2015). Este proceso viene acompañado de otro extensible a todo el ámbito científico, como es la crisis de reproducibilidad que está provocando la pérdida de credibilidad de ciertas disciplinas científicas debido a que las investigaciones no pueden reproducirse o revisarse fehacientemente (Baker 2016; Ritchie 2020). Hablamos de la proliferación de estudios "opacos" (*black boxes*), en los que no se comparten los datos ni los métodos utilizados, o se utiliza software de tipo *point and click* que hace que sea prácticamente imposible reproducir y replicar los trabajos. Desde esta perspectiva, el conocimiento científico viene construyéndose, tal y como se realiza desde el punto de vista tradicional, desde argumentos de autoridad, que no dejan mayor camino que reducir los resultados de las investigaciones a la sencilla creencia de que dicho conocimiento ha sido generado mediante una praxis analítica correcta y rigurosa, y sobre un conjunto de datos bien diseñado y estructurado. Nótese que cualquier investigador en Arqueología, entre quien se encuentra el autor de este libro, ha publicado en algún momento trabajos no reproducibles, opacos, porque esta es la tendencia más común en la disciplina arqueológica, al igual que en otras (Marwick *et al.* 2017).

En este contexto, adoptar un marco de ciencia abierta en Arqueología es un aspecto más relevante, si cabe, en las investigaciones que hacen uso de la informática, caso de la Arqueología computacional. El engranaje de técnicas y procesos analíticos que normalmente articulan estas investigaciones puede esconder carencias metodológicas importantes, alteraciones intencionadas de datos o, sencillamente, decisiones analíticas que son cruciales en la obtención de cierto tipo de resultados y que, si no se exponen abiertamente, su omisión deja una puerta abierta al fraude en la generación de conocimiento. Urge, por lo tanto, cambiar la forma en la que la disciplina arqueológica y, por extensión, los arqueólogos, entendemos los datos, los métodos y la forma en la que el conocimiento arqueológico es generado, olvidando viejas mentalidades proteccionistas y tendiendo hacia prácticas de ciencia transparente, abierta y reutilizable. Se trata, en pocas palabras, de presentar las investigaciones de forma

localizable, accesible, interoperable y reutilizable, conceptos que definen los principios FAIR para la gestión y custodia de datos científicos (Wilkinson *et al.* 2016).

Si bien el panorama actual de nuestra disciplina está cambiando, y los proyectos tienden cada vez más hacia investigaciones reproducibles, en abierto, y reutilizables (Lake 2012; Marwick *et al.* 2017), lo cierto es que este cambio es más palpable en algunas especialidades y períodos históricos, como puede ser la Prehistoria, que en otros. Además, en la disciplina arqueológica de nuestro país no existe, por el momento, un debate amplio sobre cómo se deben publicar los datos para que sean reutilizables (véase al respecto el excelente trabajo de Pardo-Gordó, Cortell-Nicolau 2023). Sigue siendo frecuente encontrarnos con monográficos y tesis doctorales que adjuntan en anexos un sinfín de tablas en formato Word con estadísticas y datos arqueológicos a los que no queda más remedio que acceder manualmente.

En la actualidad existen, sin embargo, muchas opciones para realizar investigaciones accesibles, transparentes y reproducibles. Se torna necesario, entonces, instruir a las nuevas generaciones de estudiantes e investigadores del campo arqueológico e histórico en la creación de proyectos y estudios en abierto y transparentes, con métodos de trabajo y uso de datos reproducibles que pongan de relieve la importancia de compartir la información generada y facilitar su reutilización. Apremia, pues, la introducción de cambios en los planes de doctorado nacionales hacia la integración de enseñanzas en métodos de ciencia abierta, tanto en Arqueología como en la disciplina histórica en general.

6.1. Acceso abierto a las investigaciones y datos arqueológicos

El *open access* suele referirse, en lo tocante a las publicaciones científicas, a trabajos que pueden ser descargados en línea de forma gratuita, sin coste alguno para el usuario. Esto puede lograrse, comúnmente, a través de dos vías. La publicación en acceso abierto bajo pago, conocido como *gold open access*, o el almacenamiento de versiones previas al texto publicado en formato autoarchivo, conocido como *green open access*. En esta última opción, el autor de la investigación deposita el manuscrito final (normalmente la versión ya revisada, denominada *preprint*) en un repositorio público, garantizando de esta forma el acceso libre y sin coste a su trabajo.

Las casas editoriales suelen justificar los costes del *gold open access* en términos de gastos de gestión y publicación, aunque existe gran crítica al respecto por los precios elevados, en muchos casos desorbitados.[60] Aunque algunas revistas y editoriales ofrecen exenciones de pago, lo cierto es que este condicionante económico afecta especialmente a investigadores en países en desarrollo, autores de grupos tradicionalmente infrarrepresentados, investigadores que inician su carrera o, en términos generales, a disciplinas como la arqueológica o la histórica, que no suelen contar con proyectos con el suficiente presupuesto económico como para hacer frente a estos pagos (Marwick *et al.* 2017). Nótese que una investigación accesible, sin coste, suele ser más consultada (y, por lo tanto, más citada) que aquella que es accesible sólo para suscriptores (**Figura 49**).

[60] Por poner un ejemplo: publicar un artículo con la versión *gold open access* en la revista *Quaternary International* tiene un coste de unos 3.400 dólares (véase al respecto, https://www.elsevier.com/journals/quaternary-international/1040-6182/open-access-options), algo inasumible para proyectos pequeños o investigadores con financiación reducida.

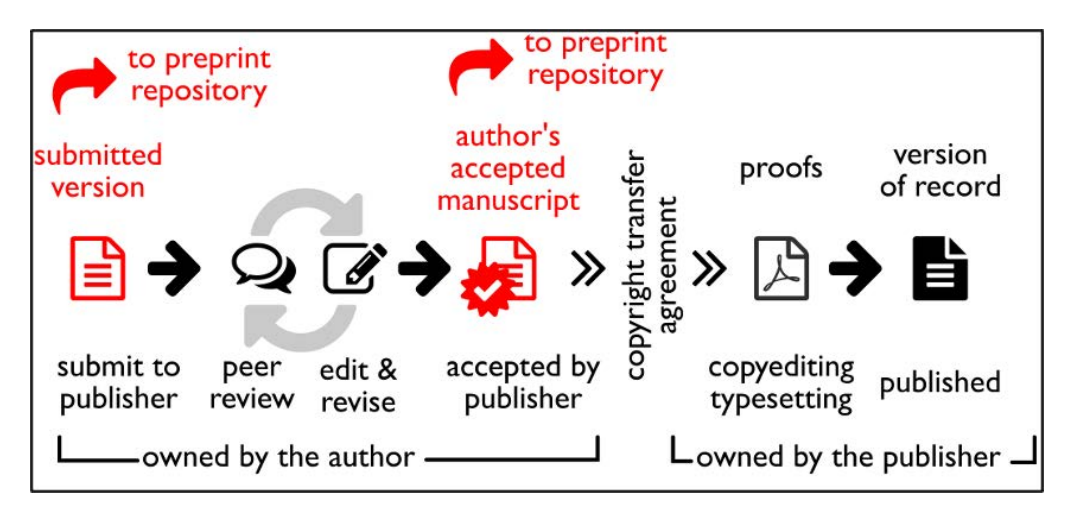


Figura 49. El ciclo de publicación en revistas académicas (obtenida de Marwick *et al.* 2017: 9).[61]

Así las cosas, en Arqueología la opción *green open access* está suponiendo, sin lugar a duda, el principal camino a seguir para hacer los manuscritos accesibles. Para ello, existen repositorios que permiten almacenar y hacer públicos los documentos en internet, algunos muy usados en humanidades y ciencias sociales, como *socarxiv*.[62] Los portales académicos *ResearchGate* o *Academia.edu*, con formato de red social pero ambos gestionados por empresas privadas con ánimo de lucro, están siendo también muy utilizados para compartir artículos, monitorizar su acceso o facilitar búsquedas relacionales.

Por otra parte, tal y como reconoce B. Marwick (*et al.* 2017: 19), el reto principal en el acceso abierto se refiere a la cesión de los datos primarios. Aparte de que es necesario considerar la propiedad intelectual de todos los datos que van a cederse con la publicación de una determinada investigación (por ejemplo, cuando se han obtenido a partir de bases de datos institucionales, como catálogos patrimoniales autonómicos), también hay que reflexionar acerca del posible impacto negativo de abrir los datos en acceso público. En el campo arqueológico ceder la localización en acceso público de determinados yacimientos puede abrir una puerta a actividades ilegales de búsqueda, detección, expolio, o incluso de destrucción del patrimonio arqueológico. Potenciales soluciones en casos sensibles pueden ser restringir el acceso, diseñar estrategias de descarga bajo petición de datos exhaustivos o limitar la precisión de la información descargable. Lo cierto es que, dependiendo del proyecto, la cantidad de datos y su precisión, será necesario realizar una sesuda discusión de la cesión de la información.

Una buena forma de hacer los datos accesibles es almacenándolos online en un repositorio público específico para Arqueología. En este sentido, el *Archaeology Data Service* de Reino Unido es uno de los más conocidos, con más de 20 años de experiencia en soporte a la investigación,

[61] Una versión *preprint* del artículo puede consultarse en: https://docs.google.com/document/d/1zuR2bv0pV23CsUhkX2DvUTa9e4x7NLfEltg9I3EkLfU/edit

[62] https://osf.io/preprints/socarxiv/

almacenamiento de datos y promoción de soluciones de acceso libre a los mismos.[63] Otros muy utilizados en la actualidad son *Zenodo*[64] y *Figshare*.[65] Recordemos que quizás la parte más importante en la elección de un determinado repositorio es la licencia, que debe cubrir el acceso y uso público de los datos (como, por ejemplo, las CC-BY), y que el repositorio elegido cuente con un plan de continuidad estable que garantice su acceso continuado y preservación (Marwick *et al.* 2017: 10).

6.2. Métodos en abierto, transparentes y reproducibles

Los métodos abiertos se refieren, en términos generales, a realizar investigaciones que puedan ser consultadas, reproducidas y replicadas por cualquiera. En los últimos años, en Arqueología han proliferado diseños metodológicos y recomendaciones para el trabajo con métodos en abierto y transparentes basados en el uso de software libre (*FOSS*), que tienen por objetivo final estructurar las investigaciones en torno a criterios reproducibles (Marwick 2016).[66] Para ello, el uso de software de programación, como *R Statistics*, es muy recomendable, porque una investigación basada en un lenguaje o "código" (*scripts*[67] de programación), es fácilmente compartible y, por definición, transparente: todos los análisis que se realizan pueden leerse y reproducirse, siempre y cuando se conozca el lenguaje de programación en el que el *script* ha sido redactado. Optar por software privado del tipo *point and click*, por ejemplo, para trabajos estadísticos (como *SPSS* o *Excel*), hace que sea muy difícil compartir los métodos y describir los análisis realizados, más allá de citar que para cierto cálculo se ha utilizado una determinada técnica. En muchos casos, además, es muy difícil saber qué tipo de algoritmo está siendo utilizado por estos programas, porque esta información no suele estar disponible, aunque es crucial a la hora de exponer de forma correcta una investigación computacional.

Por otra parte, utilizar un software de programación abierto, como *R*, abre la puerta al uso de un lenguaje de documentación de la investigación en formato libre, como puede ser *markdown* (en el caso de *R*, *RMarkdown*).[68] Como veremos posteriormente, utilizar *RMarkdown* permite documentar una investigación de forma estructurada, hacerla transparente y reutilizable por cualquiera.

El proceso de aprendizaje del lenguaje de programación *R* no es sencillo, y no existe una solución rápida. Se necesita paciencia, tiempo y sobre todo práctica, como con cualquier otro lenguaje, sea de programación o no. Hay que ser conscientes, además, de que el trabajo colaborativo suele aportar la solución más rápida y fiable a los problemas que uno se va encontrando. Por ejemplo, para la resolución de dudas concretas es frecuente acudir a plataformas conocidas como *Stack Exchange*, en las cuales se pueden formular preguntas específicas que son contestadas por una comunidad de usuarios especializados. Así, tenemos,

---

[63] https://archaeologydataservice.ac.uk/
[64] https://zenodo.org/
[65] https://figshare.com/
[66] https://rawgit.com/benmarwick/Marwick-Berlin-R-users-2017/master/Marwick-Berlin-R-users-2017.html#1
[67] Un *script* es un archivo que contiene instrucciones para ejecutar determinadas órdenes en un ordenador, como por ejemplo cierto tipo de análisis, etc. Estas instrucciones pueden escribirse utilizando cualquier lenguaje de programación, como *Python* o *R Statistics*.
[68] https://bookdown.org/yihui/rmarkdown/

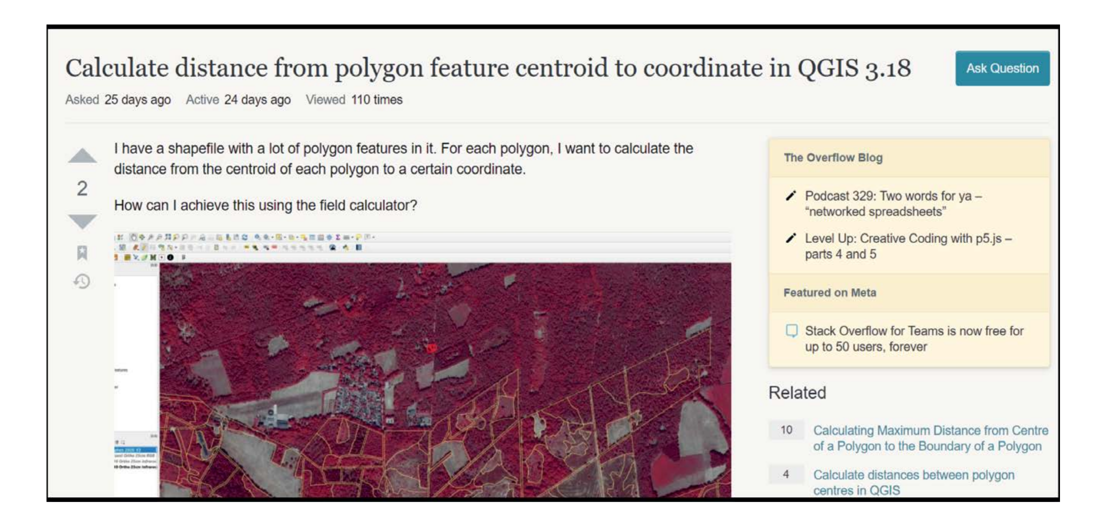


Figura 50. Ejemplo de una consulta realizada en *GIS Stack Exchange.*

por ejemplo, *GIS Stack Exchange*,[69] página orientada a dudas sobre Sistemas de Información Geográfica, o *Stack Overflow*,[70] orientada a lenguajes de programación (**Figura 50**).

Por lo que respecta a la formación en R, existen muchos cursos de introducción y métodos avanzados para análisis y manejo de datos con este lenguaje que pueden realizarse bajo pago, por ejemplo, a través de plataformas muy conocidas como *SuperDataScience*,[71] *Udemy*[72] o *Cursera*,[73]. No obstante, *R Statistics* es, por definición, un lenguaje de programación libre y de acceso abierto. Como tal, existen muchos materiales accesibles online sin coste alguno.

Una buena forma de empezar puede ser mediante la realización de los cursos ofrecidos por el propio equipo de *R*, *R Studio*[74] u otros, como los de la *Carpentry Foundation*,[75] que cuentan con materiales muy completos.[76] Podemos acudir también a proyectos colaborativos como el *Big Book of R*,[77] el *R Graphics Cookbook: Practical Recipes for Visualizing Data* (Chang 2013)[78] o, sin lugar a dudas, el libro *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data* (Wickham, Grolemund 2017).[79] Otros recomendables son los títulos de las series de la editorial

[69] https://gis.stackexchange.com/
[70] https://es.stackoverflow.com/
[71] https://www.superdatascience.com/
[72] https://www.udemy.com/
[73] https://es.coursera.org/
[74] https://education.rstudio.com/
[75] https://swcarpentry.github.io/r-novice-inflammation/
[76] Una forma interesante para buscar materiales sobre las diferentes aplicaciones y desarrollos de *R Statistics* en ciencia de datos es acudir al *Data Science with R: A Resource Compendium*, de M. Monkam, que incluye referencias a materiales y libros clasificados según sus aplicaciones en *R*: https://bookdown.org/martin_monkman/DataScienceResources_book/
[77] https://www.bigbookofr.com/index.html
[78] http://www.cookbook-r.com/; https://r-graphics.org/
[79] https://r4ds.had.co.nz/index.html; La versión de este manual en español puede encontrarse en https://es.r4ds.hadley.nz/index.html

Springer,[80] *Use R!* para temas diversos, y *Quantitative Archaeology and Archaeological Modelling* para la disciplina arqueológica.[81]

Todos estos monográficos pueden ser adquiridos en formato papel, bajo pago, pero la mayor parte cuentan con una versión online libre, en un formato denominado *bookdown*.[82] Se trata de una edición adaptada a los estándares de lenguaje *R Markdown* para su publicación en abierto, lo que permite su consulta vía web sin coste alguno.

Para cuestiones de visualización de datos es interesante combinar el manejo de las referencias anteriores con los principales monográficos sobre teoría de la visualización de gráficos, como *The Visual Display of Quantitative Information* (Tufte 2001), o *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures* (Wilke 2019), este último también con versión *bookwdown*.[83]

Por otra parte, existen materiales que diferentes investigadores han puesto accesibles de forma gratuita y online, como el pequeño workshop en visualización de datos diseñado por S. Altman, B. Behrman,[84] por citar un ejemplo, que valen para introducir muchos de los conceptos que después pueden ampliarse en los manuales anteriores.

Para el formato de sintaxis *R Markdown*, cualquiera de los trabajos citados ofrece claves para entender su funcionamiento y partes principales, con la creación de documentos, presentaciones dinámicas e informes que pueden ser exportados de forma sencilla a HTML, PDF, o Word. La referencia más completa es, quizás, el libro *R Markdown: The Definitive Guide* (Xie *et al.* 2018), que cuenta con versión *bookdown*[85] a través de la cual puede realizarse un aprendizaje completo del formato de redacción y sintaxis con *R Markdown*. Sobre *bookdown*, específicamente, puede consultarse el libro *Bookdown. Authoring Books and Technical Documents with R Markdown* (Xie 2016).

Utilizar *R Statistics* y *R Markdown* nos lleva, directamente, a hablar del uso de sistemas de control de versión o *track changes* en las investigaciones, creados para impulsar la colaboración entre diferentes usuarios, caso de Git.[86] El uso de plataformas como, por ejemplo, *Gitlab*[87] o *Github*,[88] permiten gestionar contenido analítico en colaboración, trabajar en línea y llevar un historial de los cambios realizados, con lo que se preserva la autoría inicial de las investigaciones. Además, el almacenamiento de los datos en repositorios como los ya citados *Zenodo* o *Figshare* aseguran el acceso público continuado a la información, al tiempo que mantienen la autoría inicial de los datos mediante la generación de DOIs, que suelen combinarse con licencias de uso amplio, como la MIT[89] para código (McKiernan *et al.* 2016), o las CC-BY para los datos, que permiten la reutilización siempre y cuando se atribuya la autoría inicial.

[80] https://www.springer.com/series/6991?detailsPage=titles
[81] https://www.springer.com/series/14342
[82] Un listado de muchos de los monográficos publicados en formato *bookdown*, puede encontrarse en: https://bookdown.org/home/archive/
[83] https://clauswilke.com/dataviz/
[84] https://dcl-data-vis.stanford.edu/
[85] https://bookdown.org/yihui/rmarkdown/
[86] https://git-scm.com/
[87] https://about.gitlab.com/
[88] https://github.com/
[89] https://opensource.org/licenses/MIT

6.3. Una apuesta por el uso de *R Statistics* en Arqueología

Por lo que respecta a disciplina arqueológica, en la actualidad estamos asistiendo a un despegue importante de proyectos y estudios que se estructuran en torno a *R Statistics* y otros lenguajes de programación, caso de *Python*. Esto es papable, por ejemplo, en el creciente número de publicaciones que utilizan este lenguaje (Schmidt, Marwick 2020), aunque un panorama proclive dista mucho de ser la norma común, y suele ser la acción de ciertos grupos de investigación o investigadores individuales quienes siguen marcando el camino.

En los últimos años asistimos a la publicación de monográficos que utilizan *R Statistics* en casos de estudio arqueológicos, como la adaptación a *R* del manual *Statistics for Archaeologists* de R. Drennan (2009). Otra referencia importante es el *Basic Statistical Graphics for Archaeology with R: Life Beyond Excel* (Baxter, Cool 2016), o el más reciente, *Quantitative Methods in Archaeology Using R* (Carlson 2017).

Por otra parte, es muy recomendable consultar también dos monográficos en formato *bookdown* (libre, sin coste) que, gracias al trabajo colaborativo de diferentes especialistas, proporcionan claves para realizar ciencia arqueológica con *R Statistics*. Es el caso del *Archaeological Science with R*, de B. Marwick, actualmente en elaboración, pero ya disponible para consulta y uso en *bookdown*,[90] un trabajo que se convertirá, sin lugar a duda, en un básico sobre la aplicación de *R* en Arqueología.

Este manual puede complementarse con otro, también con versión *bookdown* accesible, *How To Do Archaeological Science Using R*,[91] editado por B. Marwick y en el que diferentes especialistas exponen procesos metodológicos con análisis espaciales, modelado de base agente, cómo utilizar *R* como un SIG, realizar análisis predictivos sobre diferentes casos arqueológicos, etc.

En este sentido, son muchos los arqueólogos que en la actualidad redactan paquetes de *R Statistics* orientados específicamente para Arqueología. Uno de los compendios más completos es *CRAN Task View: Archaeological Science*, de B. Marwick,[92] en el que podemos encontrar una clasificación de paquetes por métodos y temas. Este listado puede complementarse con la consulta del *open-archeo*,[93] mantenido por Z. Batist, un compendio de software y recursos informáticos de código abierto para arqueología.

La acción de investigadores individuales, como todos los citados en este capítulo, coincide también con el trabajo de grupos más institucionalizados que cuentan con una tradición pionera en la aplicación de software libre en arqueología, nombradamente *R*. Un caso que podemos citar es el grupo ISAAK (*Initiative for Statistical Analysis in Archaeology Kiel*)[94] de la Universidad de Kiel (Alemania), cuyos miembros suelen utilizar marcos metodológicos de ciencia abierta en sus proyectos de investigación (véase por ejemplo Nakoinz, Knitter 2016); otro también reseñable es el *GIS Laboratory* del Instituto de Arqueología de la Universidad College de Londres, con el trabajo pionero de los profesores A. Bevan y M. Lake, el *Center for*

---

[90] http://benmarwick.github.io/aswr/
[91] https://benmarwick.github.io/How-To-Do-Archaeological-Science-Using-R/
[92] https://github.com/benmarwick/ctv-archaeology
[93] https://open-archaeo.info/
[94] https://isaakiel.github.io/

Social Dynamics and Complexity, de la Universidad Estatal de Arizona, dirigido por M. Barton, o el *Computational and Digital Archaeology Laboratory*, de la Universidad de Cambridge, liderado por E. R. Crema, por citar los que son más cercanos al autor.

En cuestiones de ciencia arqueológica en abierto, entre otros es el esfuerzo de B. Marwick, profesor de la Universidad de Washington, el que está suponiendo un punto de inflexión importante en nuestro campo, tanto por su trabajo en la creación de guías de buenas prácticas en ciencia arqueológica en abierto, transparente y reproducible, como su difusión a través de charlas, workshops o conferencias, cuyos materiales siempre están disponibles online para consulta y reutilización.[95]

[95] https://github.com/benmarwick

Capítulo 7.

Caso práctico. Modelos de procesos de puntos para el estudio del Megalitismo de A Serra do Barbanza (Noroeste de la Península Ibérica)

El uso de la simulación estadística está suponiendo un importante avance en los estudios de Arqueología Computacional (Rubio-Campillo 2017) porque abre la posibilidad de generar, en el caso investigaciones sobre territorios y paisajes pretéritos, modelos explícitos en los que se compara el patrón observado de los sitios arqueológicos y un conjunto de muestras aleatorias que pueden ser generadas bajo parámetros locacionales específicos. Ya hemos visto que las simulaciones aleatorias pueden crearse bajo la influencia, por ejemplo, de efectos de primer orden (variables medioambientales), condiciones aleatorias completas (CSR) o manteniendo relaciones de dependencia espacial, bien de agrupación, bien de dispersión, emulando fenómenos conocidos como eclosión, expansión o recesión (Vanzetti *et al.* 2010; Bevan *et al.* 2013; Carrero-Pazos *et al.* 2019).

Aunque sea común que estos modelos se desarrollen bajo condiciones de incerteza cronológica, pudiendo utilizar el razonamiento probabilístico para analizarlos (Crema 2015), existe un aspecto que ahora nos interesa más y que, tal y como reconocen S. Eve y E. R. Crema (2013: 272), continúa bastante inexplorado. Se trata de la posibilidad de construir y comparar, formalmente, diferentes modelos de puntos, creados con el objetivo de verificar hipótesis locacionales concretas.

Los modelos de puntos pueden ser analizados en clave de competición entre ellos gracias al uso de criterios de información (como, por ejemplo, el ya mencionado Akaike), medidas estadísticas que estudian formalmente su calidad y permiten discernir cuál es el que mejor se ajusta a los datos arqueológicos, con lo que existe la posibilidad de discutir y contrastar diferentes modelos entre sí. Este será el aspecto que ocupará nuestra atención en las siguientes líneas, utilizando el estudio de los patrones de localización de uno de los conjuntos megalíticos más señeros del Noroeste español, como es el de la Península del Barbanza.

La investigación que exponemos a continuación ha sido realizada utilizando en su totalidad software libre, nombradamente *R Statistics* para la modelización estadística, y los softwares

GRASS GIS, Quantum GIS y SAGA GIS para la construcción de las covariables espaciales. Esto nos permite compartir tanto la información arqueológica utilizada como los métodos y técnicas que hemos desarrollado, disponibles en un repositorio público (https://doi.org/10.5281/zenodo.7669911) para su consulta, descarga y uso. Asimismo, pueden introducirse cambios colaborativos en el código y comentarios acerca del mismo utilizando la plataforma Github (https://github.com/MCarreroPazos/ArqueologiaComputacionalTerritorio).

De esta forma, tanto los datos arqueológicos, el desarrollo de los métodos escogidos, así como los resultados obtenidos son accesibles públicamente y pueden ser reutilizados siempre y cuando se indique la autoría inicial, al estar consignados bajo licencia CC-BY4.0 y MIT.

7.1. El fenómeno tumular y megalítico del Barbanza

La península del Barbanza se localiza en el oeste de Galicia, y cuenta con una de las agrupaciones de monumentos megalíticos más prominentes del Noroeste español, albergando, en su parte montañosa conocida como " As chans do Barbanza", una treintena de sitios arqueológicos megalíticos, con cámaras dolménicas y túmulos térreos que han recibido particular atención por parte de la investigación gallega ya desde antiguo (véase, por ejemplo, Criado Boado *et al.* 1986; Criado Boado, Villoch Vázquez 2000; Barbeito Pose *et al.* 2015; Criado Boado 2017) (**Figura 51**).

En momentos más recientes, los estudios sobre la localización de estos monumentos han incluido la aplicación de tecnologías de información geográfica, como los SIG, estadística espacial o modelos basados en agentes (véase, entre otros, Llobera 2015; Rodríguez Rellán, Fábregas Valcarce 2015, 2019, 2023; Bustelo Abuín *et al.* 2018; Carrero-Pazos, 2023; Carrero-Pazos *et al.* 2020).

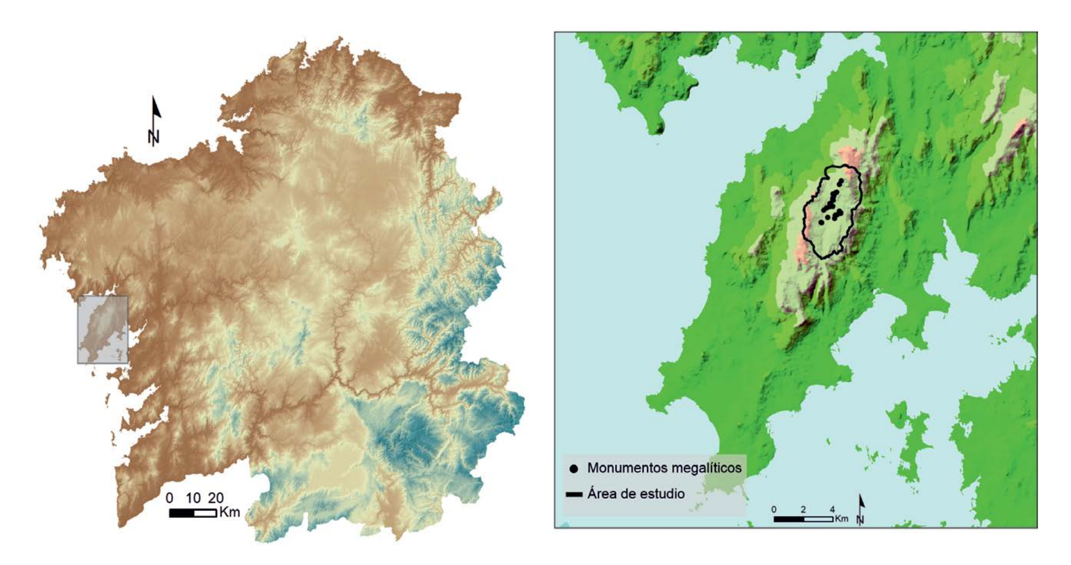


Figura 51. Izquierda: Localización de la península del Barbanza en Galicia (Noroeste de la Península Ibérica). **Derecha:** Conjuntos tumulares y dolménicos de "As Chans", localizados en las penillanuras elevadas de la sierra.

Pese a la elevada densidad de investigaciones en la zona, no existen dataciones directas para la mayor parte de los monumentos. La fecha de C14 para Casota do Páramo (CAMS-77427, 4720±120BP), obtenida a partir de pintura conservada en uno de sus ortostatos (Steelman *et al.* 2005; Carrera Ramírez 2011), indica un uso *circa* 3760-3099 cal a.C., que coincide plenamente con el discurso cronológico regional del Megalitismo del Noroeste peninsular, cuyo *summum* parece haber tenido lugar entre 3500-2700 cal a.C. Aunque dicha datación sea coherente con el contexto del Noroeste, y podamos extrapolar un uso de este espacio funerario hacia mediados del IV milenio, lo cierto es que la inexistencia de dataciones directas para la práctica totalidad de los monumentos considerados impide analizar la evolución temporal del conjunto tumular, a la vez que obliga a considerar de forma sincrónica a un conjunto de sitios arqueológicos que, con elevada probabilidad, fueron construidos y reutilizados durante arcos temporales amplios. Este hecho encuentra mayores argumentos si consideramos investigaciones en cementerios de túmulos peninsulares y europeos que han sido bien fechados, en las que se concluye que este tipo de espacios tuvieron usos recurrentes que incluso pueden pasar del milenio (por ejemplo, Jorge 1982; Makarowicz *et al.* 2018).

Los datos arqueológicos aquí utilizados forman parte de la cartografía del Megalitismo gallego creada por el grupo de estudios megalíticos de la Universidad de Santiago de Compostela, dirigido por el profesor Antón A. Rodríguez Casal (puede consultarse, al respecto, Carrero-Pazos 2017).

El trabajo de campo, junto con la modelización a través de los SIG, ha permitido plantear la importancia de todo un conjunto variables en la localización de los monumentos destacando, por ejemplo, la proximidad a rutas de tránsito por el paisaje, ejes de cuenca hidrográfica o zonas de prominencia visual (Llobera 2015; Rodríguez Rellán, Fábregas Valcarce 2015, 2019; Bustelo Abuín *et al.* 2018; Carrero-Pazos *et al.* 2020) (**Figura 52**).

Las variables definidas en el trabajo citado fueron estudiadas, también, con el objetivo de entender los efectos de segundo orden. El *modus operandi* es el ya explicado en capítulos anteriores: control de los efectos de primer orden a través de un modelo de regresión generalizado y análisis de efectos de segundo orden mediante funciones K (y derivadas) ajustadas a patrones estadísticos conocidos. Resulta interesante indicar que las funciones de correlación par, tanto las que controlan los efectos de primer orden como las que dan cuenta de los efectos de segundo orden, permitieron observar la existencia de tendencias espaciales muy agrupadas a escalas locales. Al igual que en otros casos ya explicados, las concentraciones megalíticas de As Chans do Barbanza parecen diseñarse teniendo en cuenta monumentos preexistentes, es decir, una persistencia en el enterramiento que señala que estaríamos ante lo que se ha definido como cementerios sempiternos (Fábregas Valcarce, Vilaseco Vázquez 2015: 68).

La cuestión que nos importa ahora es poner de relieve cómo la construcción de modelos de patrones de puntos más explícitos nos permite seguir testando la importancia de ciertas variables en la localización de estos sitios arqueológicos. Es decir, no llega con determinar la relevancia de una variable geográfica específica con respecto a la tendencia espacial de una distribución de sitios arqueológicos. Debemos profundizar en su entendimiento y análisis para seguir proponiendo modelos más exactos (Eve, Crema 2014: 272).

Sabemos que la visibilidad de estos monumentos fue un criterio importante a la hora de escoger su localización (cfr. Carrero-Pazos 2021). Esta visibilidad puede adquirir múltiples



Figura 52. Monumentos megalíticos de As Chans de Barbanza **(arriba),** y vistas desde lo alto de la sierra hacia la ría de Muros y Noia **(abajo)**.

tonos, significados y significantes, muchos de ellos de difícil y compleja aproximación con un SIG (véase al respecto Fábrega-Álvarez 2017; Fábrega-Álvarez, Parcero Oubiña 2019). Tomando la simplicidad como norma ahora, con objetivos didácticos diremos que esta variable puede ser entendida en términos cualitativos o cuantitativos, por ejemplo, a través del estudio de la presencia visual del túmulo en el paisaje, con relación a rutas de tránsito, zonas de prominencia topográfica, o ejes de ruptura de cuenca hidrográfica.

Teniendo como base estas explicaciones, hemos definido, a continuación, dos modelos de hipótesis simples que buscan continuar analizando la visibilidad como criterio locacional. Téngase en cuenta que el estudio de visibilidades con SIG esconde gran complejidad analítica y ahora estamos simplificando, con el objetivo de ejemplificar la creación de modelos de procesos de puntos.

Modelo hipotético 1: Visibilidad amplia. El túmulo estructura el paisaje. Si la visibilidad fue el criterio de localización más importante, podemos suponer que los túmulos se emplazarían, entonces, en zonas del paisaje prominentes, aquellas que son más visibles. Estas zonas pueden definirse desde un punto de vista cartográfico a través la construcción de paisajes visuales o *total viewshed*.

Por otra parte, la localización de estos monumentos cerca de áreas que, de forma natural, atraen la vista, como los ejes de ruptura de cuenca, permite observar cómo las sociedades neolíticas utilizaban estas arquitecturas no sólo con fines funerarios. De hecho, la presencia visual continuada de estos elementos funerarios en los paisajes ha sido interpretada como un elemento de recuerdo permanente a los muertos. Hablamos de la presencia inmanente de la tradición, que podría haber sido utilizada para mantener cierto tipo de orden social (Saxe 1970; Chapman 1981; Hanks 2008).

Modelo hipotético 2: Visibilidad restringida hacia (y desde) zonas específicas del paisaje. El túmulo megalítico como un posible marcador territorial. Este modelo analiza si la visibilidad hacia los túmulos megalíticos es restringida, o se encuentra estructurada con relación a ciertos sectores del paisaje. Nos referimos a zonas de desplazamiento natural como rutas de tránsito, áreas de ruptura de cuenca hidrográfica o lugares del paisaje que marcan los límites visuales, como las líneas de horizonte. Una presencia visual continuada del túmulo en estas localizaciones específicas ha sido un criterio interpretado, tradicionalmente, como un elemento de demarcación territorial que podría estar hablando, entre otras cosas, de aspectos del control del territorio o posesión de la tierra por parte de las sociedades neolíticas (Renfrew 1976; Thrane 1998: 275; Lagerås 2002; Tilley 2004: 197; Last 2007).

Habrá notado el lector que ambos modelos buscan responder a una misma pregunta arqueológica. ¿Pudieron actuar los túmulos megalíticos de As Chans do Barbanza como hitos en sus paisajes circundantes? El modelo 1 hace hincapié en una visibilidad general, que podemos entender ahora en términos estáticos o topográficos, mientras que el modelo 2 estudia la visibilidad como un factor restringido a ciertas zonas del territorio, en concreto, aquellas áreas relacionadas con el movimiento pedestre. Veamos, a continuación, cómo podemos factorizar, en variables susceptibles de ser analizadas con un SIG, ambos modelos hipotéticos.

7.2. Construcción de las covariables

Los factores locacionales definidos a continuación fueron creados utilizando un Modelo Digital del Terreno de 5 metros derivado de datos LiDAR, disponible para descarga en el Centro Nacional de Información Geográfica. Esta resolución nos permite contar con un detalle lo suficientemente grande como para que los resultados sean precisos, al tiempo que se minimiza el impacto de estructuras artificiales contemporáneas en los cálculos.

Para estudiar la presencia visual del túmulo en el territorio utilizaremos dos tipos de variables. En primer lugar, un mapa de visibilidad total del paisaje, a partir de los desarrollos ya comentados de M. Llobera (*et al.* 2010). Se trata de realizar un análisis de visibilidad acumulativa para todas las celdas del MDT, representando, de esta forma, desde cuantas posiciones es visible cada celda teniendo en cuenta todas las celdas restantes. Este cálculo puede realizarse en Quantum GIS con el plugin *visibility index* (Čučković 2016), aunque también es recomendable GRASS GIS y su módulo *r.viewshed.cva*[96] (**Figura 53A**).

En segundo lugar, para el cálculo de prominencia topográfica, entendiendo esta como una función de la altura diferencial entre un individuo y el entorno tal y como es percibida desde

[96] https://grass.osgeo.org/grass78/manuals/addons/r.viewshed.cva.html

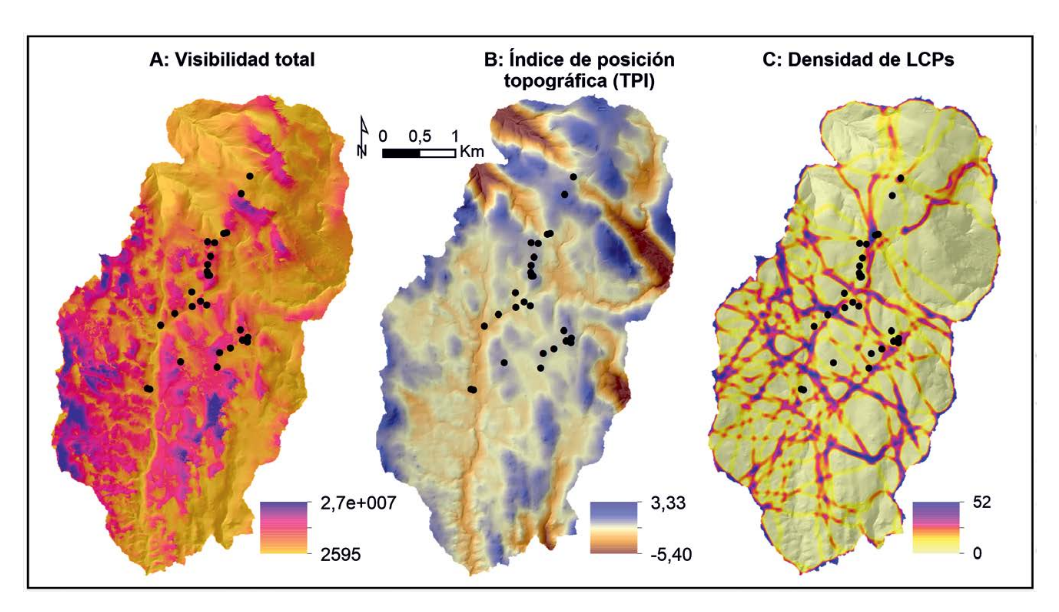


Figura 53. Variables construidas para el modelo 1. **A:** Visibilidad total del paisaje; **B:** Índice de posición topográfica (r = 500m); **C:** Densidad de rutas de menor coste.

el punto de vista del propio individuo (Llobera 2001), utilizamos SAGA GIS y su módulo *Topographic Prominence Index* (TPI). El TPI mide la posición relativa de un punto central como la diferencia entre este punto y la elevación media dentro de un radio determinado (**Figura 53B**).

Por otra parte, la creación del modelo 2 requiere la construcción de las siguientes variables: el tránsito potencial por el territorio (y la visibilidad existente desde el mismo); el cálculo de los ejes de ruptura de cuenca (la distancia a las mismas) y el cálculo de las líneas de horizonte en el paisaje. Note el lector que ninguna de las variables citadas toma como sujeto de análisis el propio monumento funerario. Esto es así porque partimos de la premisa de que los túmulos son monumentos que se articulan en un discurso natural preexistente, forman parte del paisaje, lo estructuran y moldean, con lo que nuestro primer objetivo es describir las tendencias observadas desde el punto de vista territorial para someter a comparación, posteriormente, la realidad arqueológica.

Para el cálculo del tránsito potencial por el terreno podemos utilizar las aproximaciones descritas anteriormente. Ya hemos visto que una de las más frecuentes es construir una malla de caminos óptimos que, sumados, reflejan la probabilidad de tránsito en una determinada zona. En capítulos anteriores explicamos también el uso de GRASS GIS y sus diferentes módulos para realizar este cálculo. Ahora, sin embargo, proponemos el uso de *R Statistics* y el paquete *Least Cost Path* creado por J. Lewis (2020), que nos permite agilizar y controlar el proceso de cálculo con los principales algoritmos y sus modificaciones.

Recordemos que para el cálculo de las rutas óptimas necesitamos crear, en primer lugar, una superficie que muestre el coste de tránsito. Esta superficie de coste, o fricción, fue construida, en aras de la simplicidad, utilizando simplemente dos variables que influyen en el tiempo de

demora al caminar, como son la pendiente y altitud. Un trabajo de investigación más profundo debería tener en cuenta otros factores, como la cobertura paleovegetal, etc.

El siguiente paso se concreta en la aplicación de una función de coste que calcule la relación entre el esfuerzo necesario y la demora tomada en la ruta pedestre. Existen bastantes opciones, aunque en este momento hemos optado por la función de Llobera-Slukin (Llobera, Sluckin 2007),[97] que aproxima el gasto energético metabólico al desplazarse por un paisaje en KJ/(m*kg).

Una vez tenemos el raster de fricción y la función de coste seleccionada, la librería de *R leastcostpath* nos permite seleccionar qué tipo de aproximación queremos utilizar para construir las rutas de camino óptimo. Lo interesante de utilizar este paquete, creado por un arqueólogo teniendo en consideración las necesidades frecuentes de los proyectos arqueológicos, es que gran parte del trabajo puede realizarse de forma sencilla con tan sólo unas líneas de código de programación. Aparte de permitirnos crear una ruta óptima entre dos puntos (origen A - destino B), podemos utilizar también aproximaciones más avanzadas que, de otra forma, habrían de ejecutarse de forma manual y requerirían una inversión de tiempo elevada. Nos referimos, por ejemplo, a cálculos como los de coste acumulativo desde localizaciones radiales (basado en Verhagen 2013), corredores de coste, creación de rutas de coste probabilísticas o, lo que nos interesa ahora, calcular rutas óptimas desde todas las localizaciones al resto. Un análisis que, frecuentemente, se denomina FETE (*From Everywhere to Everywhere approach*); método desarrollado originalmente por D. White y S. Barber (2012).

El FETE se basa crear rutas óptimas o de menor coste tomando como origen cada una de las celdas de un raster y como destino todas las demás, repitiendo el cálculo para la totalidad de las celdas del MDT. Como habrá notado el lector, este análisis requiere un elevado esfuerzo computacional, por lo que ahora utilizaremos una versión simplificada basada en los límites vectoriales del área de estudio. Para ello, el vector que identifica la zona de trabajo debe ser descompuesto en puntos lineales: por ejemplo, 1 punto cada 500m, lo que, en nuestra zona de trabajo supone un total de 54, que serán utilizados para calcular, como origen y destino, el conjunto de rutas óptimas para la totalidad del territorio. El resultado del análisis será un vector que incluye aquellas áreas del paisaje en las que existe mayor o menor número de rutas óptimas. Partiendo de él podemos generar una superficie raster que indique la densidad de las rutas óptimas, por ejemplo en GRASS GIS, con el módulo *v.kernel* **(Figura 54)**. Esta superficie de densidad de rutas óptimas fue incluida en el modelo 1, al corresponderse con la movilidad natural por el paisaje.

Finalmente, para analizar la relación visual entre el movimiento por el paisaje y los túmulos megalíticos podemos calcular la visibilidad existente desde las propias rutas de tránsito potencial. Esto nos permite analizar qué zonas del territorio pueden observarse desde las mismas y, por lo tanto, cuantificar si los monumentos son visibles desde estas partes específicas

[97] Las cuestiones son bastante más complejas. Tal y como reconocen J. Márquez Pérez (*et al.* 2015: 48-49), entre los factores implicados en el cálculo del tiempo de demora, la inclinación del terreno se erige como una de las variables clave, pues su magnitud y sentido influyen en el coste metabólico invertido: para un caminante, afrontar una pendiente positiva (ascender) supone un esfuerzo mayor que caminar sobre una superficie plana, mientras que desplazarse por una pendiente negativa (descender) requiere un esfuerzo que puede ser menor, igual o superior al de su homóloga positiva, dependiendo de su gradiente, pues a partir de determinado valor consume más energía al ir cuesta abajo que cuesta arriba (Llobera, Sluckin 2007).

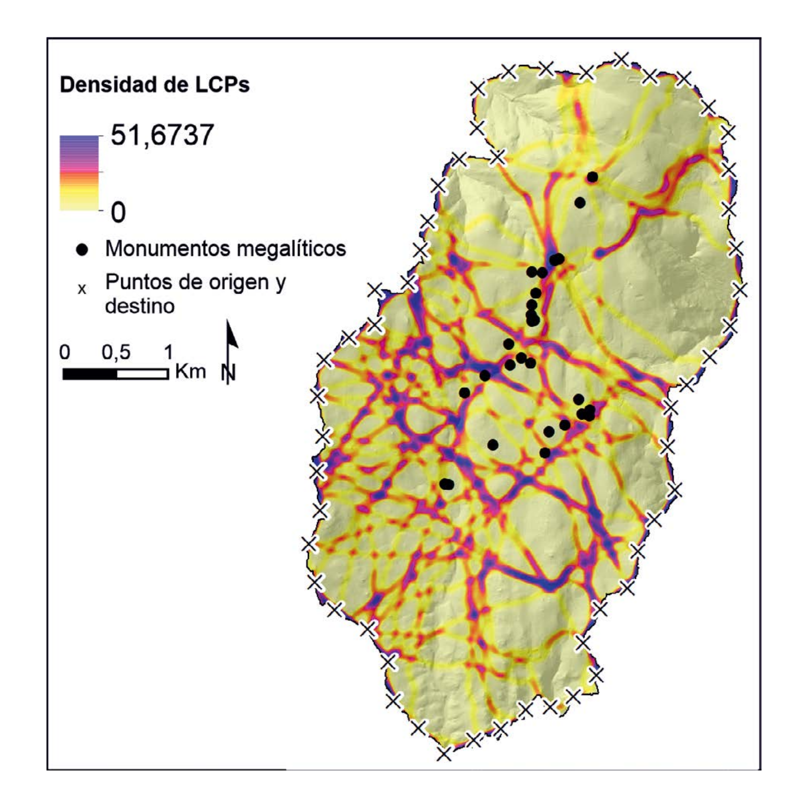


Figura 54. Densidad de rutas óptimas (*least cost paths*) para As Chans de Barbanza.

del paisaje. Su cálculo puede ser realizado, de nuevo, con la herramienta *visibility index* de Quantum GIS, utilizando como puntos de observación el vector de rutas óptimas obtenido en pasos anteriores. Para ello, recordemos que hay que descomponer este vector en puntos lineales. En este caso hemos situado 1 punto cada 100 metros, lo que supone, en total, 5.454 puntos que serán utilizados para el cálculo de visibilidad en QGIS (**Figura 55A**).

Para calcular la variable de distancia a ejes de ruptura de cuenca hidrográfica es necesario empezar por crear, primero, las cuencas hidrográficas de la zona. Esto puede hacerse de forma sencilla en GRASS GIS con el módulo *r.watershed*. A partir de aquí, hay que calcular la distancia que existe desde cualquier punto del mapa a los ejes de las cuencas. Por razones de sencillez se ha optado por la distancia euclídea (*r.grow.distance*, en GRASS GIS), aquella medida en línea recta, aunque un trabajo más oportuno requeriría de una reflexión sobre si este método es una buena forma de acercarse a la relación espacial entre localizaciones determinadas (**Figura 55B**).

La última variable definida por el modelo 2 es la densidad de líneas de horizonte, que puede calcularse también con la herramienta *visibility index* de Quantum GIS. En este caso hemos optado por realizar una aproximación a la densidad total de líneas de horizonte, es decir, se ha calculado la línea de horizonte visual para todas las celdas del MDT que, sumadas, nos proporcionan un raster de densidades de líneas de horizonte (**Figura 55C**).

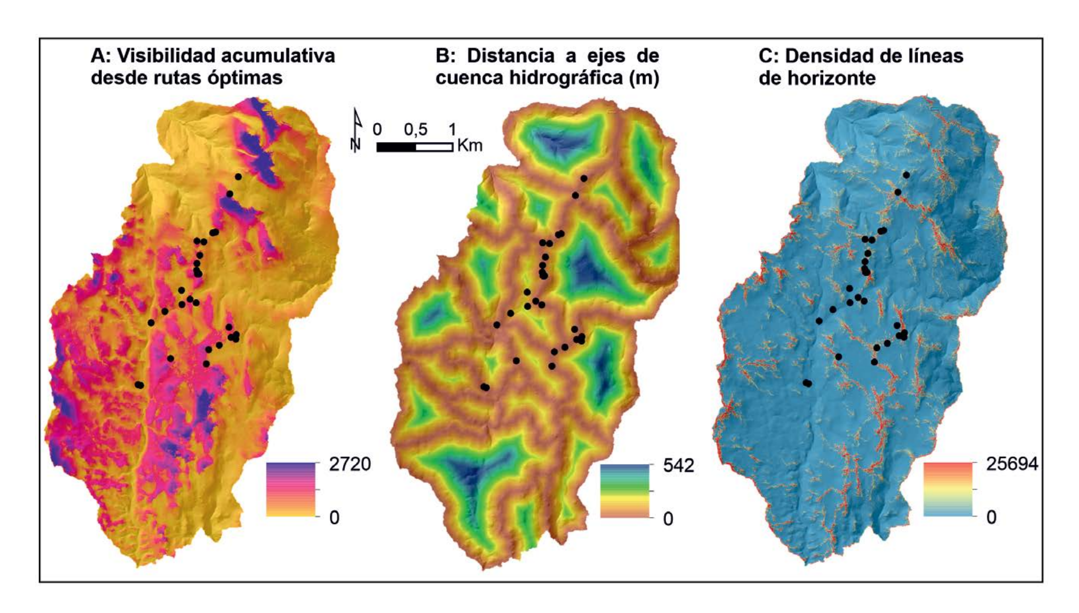


Figura 55. Variables construidas para el modelo 2. **A:** Visibilidad acumulativa desde rutas óptimas; **B:** Distancia a ejes de cuenca hidrográfica (m); **C:** Densidad de líneas de horizonte.

7.3. Formalización y discusión de los modelos

Las hipótesis locacionales definidas anteriormente, que fueron transformadas en modelos teóricos para estudiar la visibilidad de los conjuntos megalíticos del Barbanza, pueden formalizarse como eventos de un proceso de puntos para explorar estadísticamente su validez. Siguiendo la investigación de S. Eve y E. R. Crema (2013), hemos optado por formalizarlas como eventos de un proceso de puntos no homogéneo conocido como *Poisson*, en el que la intensidad del proceso es una función de la frecuencia de referencia y de los efectos adicionales inducidos por las variables externas, es decir, las variables raster modeladas.

Desde esta perspectiva, el modelo 1 asumirá, entonces, que la intensidad del proceso de puntos varía en función de la visibilidad topográfica, la prominencia del terreno y la densidad de rutas óptimas. El modelo 2, por su parte, asumirá que la intensidad del proceso de puntos varía en función de la visibilidad desde las rutas óptimas, la distancia a ejes de cuenca hidrográfica, y la densidad de líneas de horizonte.

Por último, a efectos comparativos, es necesario crear también un modelo nulo (modelo 3), en el que la intensidad del proceso de puntos es homogénea sobre el espacio e independiente a todas las variables.

La selección del mejor modelo posible puede realizarse mediante el ya explicado Criterio de Información Akaike (AIC), que nos permite escoger cuál de los modelos propuestos es más probable que haya generado el patrón observado.[98] En términos generales, ya hemos dicho

[98] Si bien el AIC es una estimación estadística muy interesante para la elección del mejor modelo posible, existen otras aproximaciones que contemplan combinaciones de variables, ponderación de la importancia de los modelos en proporción a los valores AIC, etc., aspectos que en trabajos más específicos deberían explorarse (véase al respecto Baddeley *et al.* 2015).

que cuanto menor sea el AIC más preciso será el modelo, y los resultados de los tres modelos son los que siguen: modelo 1, AIC= 827.1675; modelo 2, AIC= 766.2437; modelo 3, AIC= 829.2951. El AIC indica, por tanto, que el modelo 2 es, con el menor número de suposiciones, aquel más preciso.

Posteriormente, podemos examinar los efectos de la dependencia espacial inherente al proceso de puntos midiendo los denominados residuales de los modelos (Baddeley *et al.* 2011). Los residuales son aquellas desviaciones significativas del registro empírico con respecto a las expectativas dictadas por los modelos, es decir, aquellos casos en los que observamos una densidad de túmulos megalíticos superior a la esperada bajo determinadas escalas espaciales. Los residuales pueden utilizarse, entonces, como guía para elegir un modelo de interacción más adecuado (Eve, Crema 2013: 273).

Para medir los residuales de los modelos podemos utilizar la función K residual, que nos permite continuar entendiendo los modelos creados y su ajuste a la distribución de túmulos megalíticos (**Figura 56**).

En el caso de la **Figura 56**, las simulaciones aleatorias (bandas grises) están condicionadas por las variables que componen los modelos (excepto en el caso del modelo 3, que está creado bajo condiciones aleatorias de CSR). Desde esta lógica se infiere entonces que, si la distribución espacial de los túmulos megalíticos (línea negra en la **Figura 56**) pudiese explicarse, por ejemplo, por la dependencia espacial inducida por una visibilidad locacional amplia (definida en el modelo 1), esperaríamos entonces que la tendencia de los datos arqueológicos se localizase dentro de la banda gris aleatoria. Esto no ocurre, por lo que podemos concluir que el modelo 1 no consigue explicar la tendencia espacial de los túmulos megalíticos.

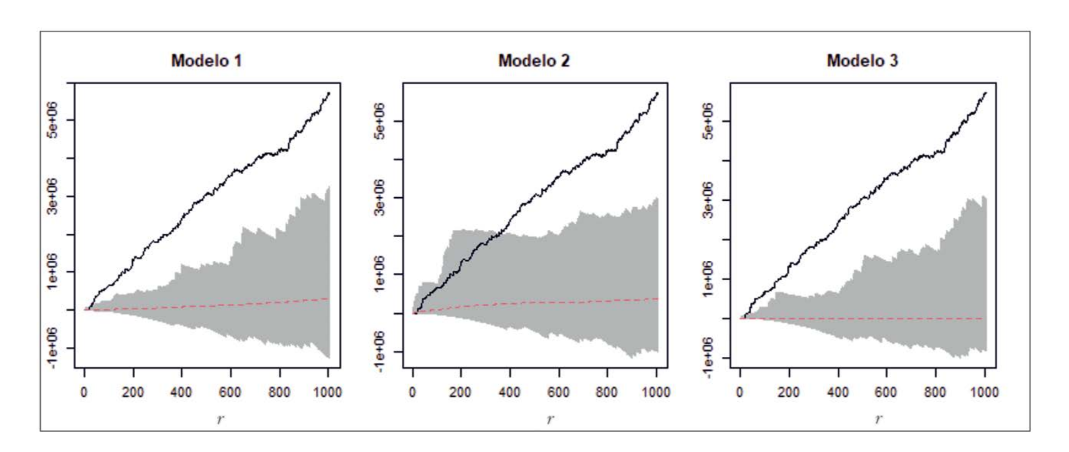


Figura 56. Gráficos de la función K residual para los diferentes modelos creados. En los modelos 1 y 2 las simulaciones aleatorias se basan en las diferentes variables construidas. El modelo 3 es un modelo nulo, puramente aleatorio. Las líneas negras del gráfico se corresponden con la distribución observada (túmulos megalíticos).

El modelo 2, en cambio, proporciona resultados más interesantes. En este caso, la distribución espacial de los túmulos megalíticos (que muestra agrupación) sí puede explicarse, aunque sólo a escalas locales, por una visibilidad restringida, aquella relacionada con el movimiento por el paisaje, porque la línea negra que identifica los monumentos está situada, hasta los 400m, sobre el conjunto de simulaciones aleatorias.

Los resultados permitirían indicar, entonces, que la visibilidad es un criterio locacional relevante a la hora de entender la distribución de los túmulos megalíticos de As Chans de Barbanza y que, con elevada probabilidad, la visibilidad orientada al movimiento por el paisaje fue uno de los criterios específicos seguidos a la hora de construir estos monumentos, sobre todo a escala local. De aquí podemos inferir que los túmulos de As Chans do Barbanza pudieron haber ejercido como hitos territoriales, al mantener una visibilidad orientada a las áreas del paisaje por las que las poblaciones se movían, aspectos ya propuestos por investigadores anteriores (por ejemplo, Criado Boado, Villoch Vázquez 2000; Rodríguez Rellán, Fábregas Valcarce 2019, 2023). Se trata de un inmanente recuerdo a los muertos; un hito territorial permanente; un uso explícito del paisaje, ahora re-estructurado en torno a elementos culturales.

Nótese que estas conclusiones están basadas en un ejemplo didáctico más que en una investigación profunda sobre el tema, pues indudablemente un trabajo más oportuno requeriría de un debate previo sobre la selección y construcción de las variables, los efectos de primer y segundo orden. Pero lo cierto es que, con las variables iniciales seleccionadas, estos son los resultados más ajustados que podemos obtener, aunque sería conveniente la inclusión de otros factores que, con total seguridad, han jugado un papel clave en la conformación del patrón de la distribución de monumentos megalíticos en la zona (véase al respecto, por ejemplo, Carrero-Pazos *et al.* 2020, Carrero-Pazos, 2023).

La metodología expuesta pone de relieve cómo los procesos de puntos permiten profundizar en la explicación de los criterios locacionales que están detrás de la formación de una determinada distribución de sitios arqueológicos y sus patrones espaciales. Estos métodos pueden ser aplicados, además, a cualquier casuística arqueológica, sean asentamientos poblacionales, monumentos funerarios, materiales arqueológicos, etc.

7.4. Compartir la investigación. Almacenamiento, publicación en línea, y lenguaje *Rmarkdown*

En capítulos precedentes indicamos la importancia de compartir, junto a la publicación de una determinada investigación, la información que lo articula, tanto los datos de base y métodos aplicados, como los resultados obtenidos. Es decir, se trata de realizar investigaciones “ejecutables” (Marwick 2016), el único camino posible hacia una democratización total del conocimiento entre la comunidad científica. Recordemos también que esto puede hacerse sólo si optamos por utilizar software libre, que por definición es transparente, de acceso abierto y gratuito.

Proporcionaremos, a continuación, un ejemplo explicativo del código *Rmarkdown* generado para la investigación anterior. Dicho trabajo está disponible en un repositorio *Github*, bajo el siguiente enlace (https://github.com/MCarreroPazos/ArqueologiaComputacionalTerritorio).

En él puede encontrarse la siguiente estructura de datos, sobre una carpeta directorio que hemos denominado *Capítulo 7*. Recomendamos su descarga y apertura para atender a las explicaciones que siguen:

Carpeta/directorio/archivo	**Datos que contiene**
csv	Archivo en formato *csv* con los datos arqueológicos (túmulos megalíticos de As Chans de Barbanza). Se incluye un número identificativo y las coordenadas espaciales (UTM, ETRS89, huso 29N)
grids	En esta carpeta pueden encontrarse todos los ráster generados mediante métodos SIG: *mdt05.tif* = modelo digital de elevaciones de 5 metros *dist_cuencas_hidro.tiff* = raster de distancia a cuencas hidrográficas *lcp_density100m.tiff* = raster de densidad de rutas de menor coste, calculado a partir de un análisis de densidades kernel con ventana de 100 metros *total_horizon.tif* = raster de densidad de líneas de horizonte *total_viewshed.tif* = raster de visibilidad total del paisaje *TPI.tif* = raster del índice de posición topográfica *visib_acum_desde_rutas.tiff* = raster de la visibilidad acumulativa calculada desde las rutas óptimas
shp	En esta carpeta pueden encontrarse los archivos *shapefile* generados: area.shp = vector del área de estudio (polígono) area_puntos_500m.shp = vector del área de estudio descompuesto en puntos situados de forma regular cada 500 metros fete_lcp.shp = output del cálculo de rutas óptimas FETE fete_lcp_puntos.shp = vector de rutas óptimas descompuesto en puntos situados regularmente cada 100 metros
log.R	Archivo *script* de *R Statistics* que contiene el código creado para ejecutar la investigación
log.Rmd	Archivo *script* de *Rmarkdown* que contiene el código generado para compartir la investigación en lenguaje *Rmarkdown*.
log.docx	Output de *Rmarkdown* para Word
log.html	Output de *Rmarkdown* para HTML

A continuación, se incluyen y comentan las partes del código que pueden encontrarse tanto en un archivo Word como HTML en el citado repositorio. Ambas opciones pueden ser directamente exportadas desde *R* utilizando la función *knit*. No se espere de lo que sigue un tutorial sobre este lenguaje de programación, sino un ejemplo de cómo una investigación ejecutable y reproducible permite realizar un trabajo totalmente transparente. La versión incluida a continuación se ha simplificado en aras de realizar una explicación más sencilla. Consúltese el repositorio para observar una versión completa.

Como *R* se basa en el uso de diferentes librerías individuales, denominadas también paquetes, lo primero que debemos hacer es cargar estos paquetes en la sesión de trabajo. Esto es lo que se realiza en la primera parte, o *set up*. Dependiendo de los análisis que vayamos a realizar, será necesario cargar una librería u otra. Desde el inicio, además, hay que establecer la carpeta directorio sobre la que trabajaremos, en la que se almacenan los datos arqueológicos y en la que los análisis serán ejecutados. En este caso la hemos denominado *Capítulo 7*.

```
Primera parte "set up"
Cargar las librerías y establecer el directorio de trabajo
spatpack<-c("raster","spatstat","rgdal","maptools", "MuMIn", "MASS")
lapply(spatpack, require, character.only=TRUE)
setwd("~/Capítulo 7") # Establecer el directorio de trabajo
```

Además, es necesario cargar la información de inicio: los datos arqueológicos y el área de estudio. En este caso, la capa de los túmulos megalíticos (en formato *csv* delimitado por comas) se encuentra en la carpeta denominada "csv", y el vector del área de estudio (en formato *shapefile*) en la carpeta "shp".

```
Cargar información de inicio (yacimientos y área de estudio)
tumulos <- read.table(file="csv/yacimientos.csv",header=TRUE, sep=";")
area_estudio <- readOGR(dsn="shp/area.shp", layer="area")
```

La segunda parte del código se ocupa de ejecutar ciertos pasos previos necesarios para crear y analizar correctamente los modelos de patrones de puntos. Se trata de crear un patrón espacial (objeto tipo *ppp*, *point pattern process*) a partir de los sitios arqueológicos y la ventana analítica (el vector del área de estudio).

```
Segunda parte: pasos previos para el análisis de patrón de puntos
Crear el patrón de puntos de los sitios arqueológicos
area <- as(area_estudio,"owin")
sppp <- ppp(x= tumulos$UMTX, y= tumulos$UMTY, window=area)
```

Posteriormente, cargamos los diferentes mapas raster generados con métodos SIG, es decir, las covariables independientes que utilizaremos para crear los modelos de procesos de puntos.

```
Cargar rasters generados con métodos SIG
lcp_dens <- raster("grids/lcp_density100m.tiff")
total_view <- raster("grids/total_viewshed.tif")
tpi <- raster("grids/TPI.tif")
dist_ejes_cuencas <- raster("grids/dist_cuencas_hidro.tiff")
total_horizon <- raster("grids/total_horizon.tif")
visib_rutas <- raster("grids/visib_acum_desde_rutas.tiff")
```

Estas variables tienen que ser convertidas a imágenes para que puedan ser incluidas en un análisis de patrón de puntos, paso que realizamos a continuación.

```
Convertir las variables a objetos espaciales (para spatstat)
lcp_densidad <- as.im(as(lcp_dens, "SpatialGridDataFrame"))
t_view <- as.im(as(total_view,"SpatialGridDataFrame"))
tprom <- as.im(as(tpi,"SpatialGridDataFrame"))
d_ejes_cuencas <- as.im(as(dist_ejes_cuencas,"SpatialGridDataFrame"))
t_horizon <- as.im(as(total_horizon,"SpatialGridDataFrame"))
visib_desde_rutas <- as.im(as(visib_rutas,"SpatialGridDataFrame"))
```

La tercera parte se concreta en la creación y ejecución de los modelos de patrones de puntos. Al ejecutarlos obtenemos las tablas de resultados en las que podemos observar la interacción de las variables y, por ejemplo, su valor *p* (columna Ztest).

Tercera parte: creación y ejecución del análisis de patrón de puntos

Modelo 1

```
(modelo1 = ppm(sppp~lcp_dens+t_view+tprom))
## Nonstationary Poisson process
##
## Log intensity:  ~lcp_dens + t_view + tprom
##
## Fitted trend coefficients:
##   (Intercept)      lcp_dens       t_view        tprom
## -1.389750e+01  2.618525e-02  8.945376e-08  2.217464e-01
##
##                Estimate         S.E.       CI95.lo       CI95.hi Ztest
## (Intercept) -1.389750e+01 3.566937e-01 -1.459661e+01 -1.319839e+01   ***
## lcp_dens     2.618525e-02 5.196848e-02 -7.567111e-02  1.280416e-01
## t_view       8.945376e-08 4.090090e-08  9.289462e-09  1.696181e-07     *
## tprom        2.217464e-01 2.459185e-01 -2.602451e-01  7.037379e-01
##                   Zval
## (Intercept) -38.9620059
## lcp_dens      0.5038679
## t_view        2.1870852
## tprom         0.9017068
```

Modelo 2

```
(modelo2 = ppm(sppp~d_ejes_cuencas+t_horizon+visib_desde_rutas))
## Nonstationary Poisson process
##
## Log intensity:  ~d_ejes_cuencas + t_horizon + visib_desde_rutas
##
## Fitted trend coefficients:
##       (Intercept)   d_ejes_cuencas        t_horizon visib_desde_rutas
##     -1.180705e+01    -2.520354e-02     3.029313e-05     -8.611419e-05
##
##                      Estimate    S.E.    CI95.lo       CI95.hi      Ztest
## (Intercept)   -1.180705e+01   3.819420e-01  -1.255564e+01  -1.105845e+01   ***
## d_ejes_cuencas-2.520354e-02   6.173410e-03  -3.730320e-02  -1.310388e-02   ***
## t_horizon     3.029313e-05   7.408012e-06  1.577369e-05   4.481256e-05   ***
## visib_desde_rutas -8.611419e-05 3.818737e-04  -8.345729e-04   6.623445e-04
##                       Zval
## (Intercept)        -30.9131972
## d_ejes_cuencas     -4.0825955
## t_horizon          4.0892382
## visib_desde_rutas  -0.2255044
```

Modelo 3

```
(modelo3 = ppm(sppp~1)) # Modelo nulo
##                        Stationary                        Poisson                        process
##                                          Intensity:                                     1.736378e-06
##                      Estimate              S.E.             CI95.lo          CI95.hi        Ztest             Zval
## log(lambda) -13.26371   0.1856953  -13.62767   -12
```

También podemos calcular los criterios de información Akaike para cada modelo.

```
Comparar los AIC's para los modelos
Modelo                    | AIC
--------------------------|-------------
Modelo 1                  | 827.1675454
Modelo 2                  | 766.2437141
Modelo 3                  | 829.2951307
```

Y finalmente, calculamos las K residuales para cada modelo, generando un gráfico común para los tres modelos.

```
Calcular las K residuales
residualK_mod1 = envelope(modelo1,Kres,correction="best",nsim=999)
residualK_mod2 = envelope(modelo2,Kres,correction="best",nsim=999)
residualK_mod3 = envelope(modelo3,Kres,correction="best",nsim=999)

Generar un gráfico de las funciones K residuales de los modelos
par(mfrow=c(1,3))
plot(residualK_mod1, legend = F, main="Modelo 1")
plot(residualK_mod2, legend = F,main="Modelo 2")
plot(residualK_mod3, legend = F,main="Modelo 3")
par(mfrow=c(1,1))
```

Como el lector habrá notado, siguiendo los pasos anteriores esta investigación es accesible, puede ser leída por cualquiera que conozca el lenguaje y puede reproducirse en cualquier ordenador, obteniendo el gráfico final de las funciones K residuales para analizar los modelos locacionales propuestos. Un ejercicio que sería interesante realizar es adaptar esta metodología, y por tanto el código de *R* aquí descrito, a otro caso de estudio. Nótese también que el repositorio en el que la investigación está almacenada indica que los datos cartográficos y arqueológicos están sujetos a licencia CC-BY4.0,[99] y el código de *R* y *R Markdown*

[99] https://creativecommons.org/licenses/by/4.0/deed.es

bajo la licencia MIT.[100] En ambos casos, el usuario interesado es libre de compartir, copiar y redistribuir el material, adaptarlo o construir material nuevo a partir del mismo, bajo la única condición obligatoria de atribuir la autoría original. En otras palabras, para utilizar este material es obligatorio citar la autoría y fuente de la que se han extraído los datos y el código, una práctica que de facto es, o debería ser, de sentido común en la investigación.

[100] https://opensource.org/licenses/MIT

Capítulo 8.

Caso práctico. Patrones de localización en el poblamiento de la Edad del Hierro de Galicia (Noroeste de la Península Ibérica)

El segundo caso de estudio se centra en el análisis de los patrones de localización del poblamiento de la Edad del Hierro en un sector del noroccidente gallego. En concreto, utilizaremos los resultados de la investigación de C. Parcero Oubiña y S. Nión Álvarez (2021) titulada *Forms of settlement inequality over space. A GIS-based method for measuring differences among settlements* para llevar a cabo un estudio espacial de un conjunto de castros localizados en el centro-norte de Galicia. El trabajo que utilizamos, que puede consultarse en su versión publicada en la revista *Journal of Archaeological Science: Reports* (bajo suscripción),[101] o la versión *preprint* disponible en Academia.edu,[102] proporciona tanto los datos arqueológicos como los *scripts* analíticos diseñados para realizar la investigación. Gracias a ello, la investigación puede ser reproducida, replicada y usada como base para otras posteriores, caso de este capítulo.

La investigación de Parcero y Nión se centra en desentrañar la existencia de desigualdades en los asentamientos castreños de un sector específico del Noroccidente gallego. En concreto, los autores diseñan varios métodos con SIG para calcular, medir y expresar cuantitativamente la desigualdad entre aproximadamente un millar de asentamientos de la Edad del Hierro. Su investigación se concreta en la realización de inferencias acerca del tamaño de las poblaciones que habitarían dichos castros y la construcción de elementos defensivos tipo terraplén. Para ello, cuantifican y comparan la extensión de los asentamientos, el porcentaje total de superficie ocupada por los terraplenes defensivos y el posible esfuerzo invertido en su construcción.

Sus conclusiones apuntan hacia la existencia de marcadores de desigualdad en torno al tamaño de los asentamientos o la monumentalidad defensiva, con identes diferencias entre áreas territoriales. En opinión de los autores, los sectores de costa parecen haber mantenido criterios de desigualdad más tradicionales, como el tamaño de los asentamientos, mientras que en áreas del interior el tamaño de los grupos humanos parece ser más equivalente, y

[101] https://www.sciencedirect.com/science/article/pii/S2352409X20305307#s0105
[102] https://www.academia.edu/45272213/Forms_of_settlement_inequality_over_space_A_GIS_based_method_for_measuring_differences_among_settlements

las diferencias entre los castros pudieron tener expresión a través de la construcción de complejos sistemas defensivos con monumentalidad diversa y desigual inversión en el esfuerzo constructivo. Ello tiene explicación, siguiendo a los autores, si dichas conclusiones se enmarcan sobre las propuestas de A. González-Ruibal (2011), para quien el poblamiento de esta zona se correspondería con un modelo de "sociedades heroicas" con liderazgos militares (Parcero Oubiña 2003).

La cuestión que queremos estudiar en este capítulo es si la localización de estos yacimientos es homogénea en el paisaje o si, por el contrario, el patrón es heterogéneo o aleatorio. ¿Están los castros localizados en partes específicas del paisaje? ¿Existe algún patrón de territorialidad entre ellos?

Este capítulo no pretende otra cosa que ejemplificar cómo la aplicación de ciertos métodos y técnicas con base SIG y análisis espacial de patrones de puntos permiten entender mejor las dinámicas de poblamiento prehistóricas, aportando base cuantitativa a los argumentos desarrollados por investigadores previos y contribuyendo a la discusión de las diferentes formas de poblamiento del espacio en el pasado.

Al igual que con el caso práctico anterior, las investigaciones propuestas en este capítulo fueron realizadas siguiendo los actuales estándares de ciencia abierta y reproducible mediante el uso de *R Statistics*. Tanto los datos arqueológicos como los diferentes *scripts* creados pueden consultarse en acceso abierto a través del repositorio Zenodo (https://doi.org/10.5281/zenodo.7669911), junto con GitHub para trabajo colaborativo, por si algún usuario interesado quisiese introducir cambios, corregir posibles errores o ampliar alguna parte, etc. (https://github.com/MCarreroPazos/ArqueologiaComputacionalTerritorio).

8.1. Introducción, área de estudio y datos arqueológicos

Este trabajo se centra en el estudio de los patrones de localización de los asentamientos tipo castro de la Edad del Hierro en un sector específico del noroccidente gallego. Durante este período, la ocupación de un territorio heterogéneo en términos topográficos, como es el Noroeste, se llevó a cabo a través de asentamientos fortificados conocidos como "castros" (Parcero Oubiña *et al.* 2017). La variabilidad en forma y tamaño es una de las características más definitorias de este tipo de poblamiento, predominando castros de pequeño tamaño en los períodos iniciales (1-2 ha, hasta los siglos II-I a.C.) hasta la proliferación de grandes *oppida* (**Figura 57**). Tal y como indican C. Parcero y S. Nión en el trabajo referido anteriormente, es durante las etapas finales de la Edad del Hierro (S. IV a.C. en adelante) cuando se documenta un aumento significativo de la densidad de estos asentamientos.

El debate en torno a las formas de organización social de estas comunidades en el Noroeste peninsular ha sido prolijo, existiendo divergencias entre quienes consideran que existió una tendencia hacia la desigualdad social a medida que avanza la Edad del Hierro (González-Ruibal 2011; Parcero Oubiña 2013) y aquellos que proponen un modelo de sociedades segmentarias continuado hasta la conquista romana (Currás, Sastre 2020). Dado que este capítulo tiene por objetivo el discurso metodológico y no tanto el interpretativo, remitimos al lector a la bibliografía referenciada para ampliar dichas discusiones.

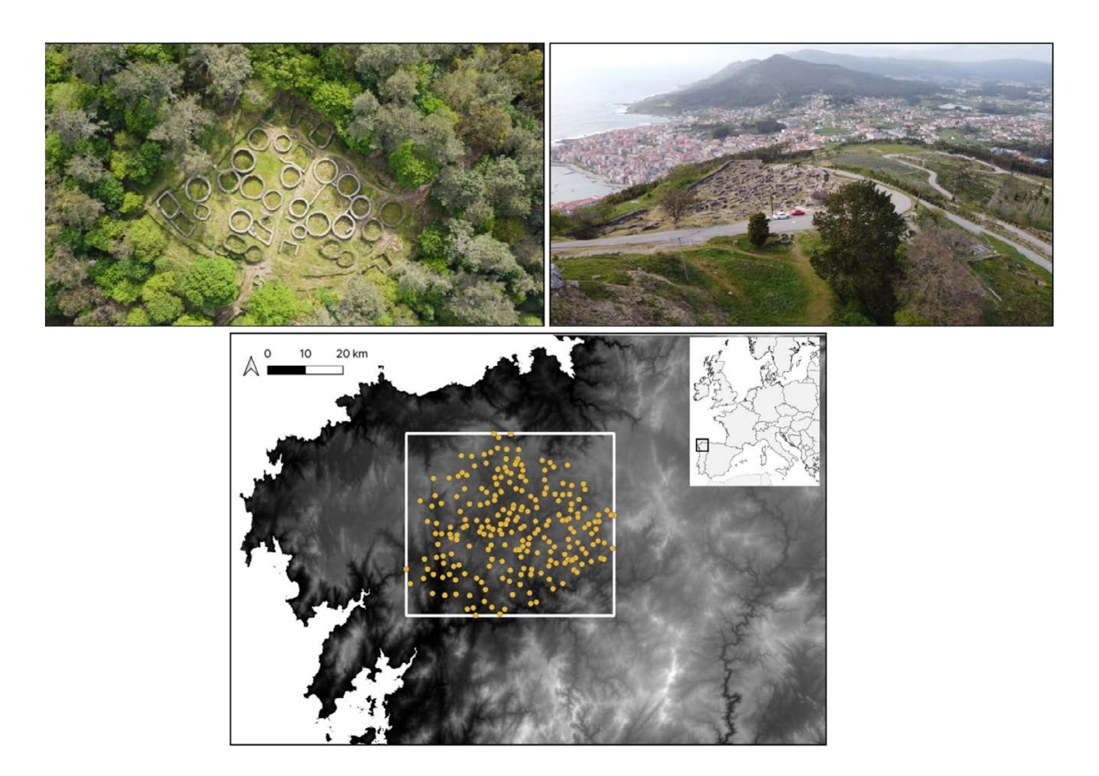


Figura 57. Arriba (izquierda): vista aérea del castro de Borneiro (Cabana de Bergantiños, A Coruña). **Arriba (derecha):** vista aérea del castro de Santa Trega (A Guarda, Pontevedra). **Abajo**: área de estudio definida en este trabajo, con los castros punteados.

El área de estudio que hemos seleccionado se localiza en el centro-oeste gallego, incluyendo las actuales comarcas de Ordes, Arzúa y Santiago de Compostela (**Figura 57, abajo**), con un total de 229 castros de diferente tamaño, planta y morfología. Tal y como hemos indicado anteriormente, en el estudio original los autores analizan un área más grande, incorporando un número más elevado de yacimientos castreños.

En términos geográficos estamos ante una zona de contrastes, con un pre-litoral gobernado por valles escarpados y zonas planas, como la depresión meridiana, y áreas montañosas dependientes de la Dorsal Meridiana Gallega, cordal montañoso que atraviesa Galicia desde el cabo de Estaca de Bares, al norte, hasta el río Miño en la frontera con Portugal, al sur.

Existe una premisa importante que es necesario aclarar desde el inicio. Este trabajo asume, por definición, estacionariedad en la muestra de datos, lo que significa que los diferentes yacimientos serán estudiados en clave de contemporaneidad. Si bien esto podría cumplirse en algunos casos, en otros sabemos con seguridad que esto no es cierto. La falta de excavaciones y programas de dataciones, la dificultad de fechar categóricamente estos asentamientos junto a que la mayor parte probablemente estuvieron ocupados durante arcos temporales amplios, provoca que no podamos sino adoptar una perspectiva global ahora. Es decir, estudiar el conjunto de castros sin tener en cuenta su variable temporal específica.

8.2. Construcción de las covariables

El primer paso a la hora de estudiar las tendencias locacionales es construir las diferentes variables que serán utilizadas posteriormente en un modelo de regresión logística, buscando seleccionar aquellas que predigan acertadamente la localización de estos asentamientos.

Tal y como adelantábamos, no existe una recomendación simple y definitiva sobre qué variables se deben estudiar, sino que éstas deben emanar de un trabajo bibliográfico, factorizando los criterios locacionales que puedan extraerse de los modelos histórico-arqueológicos sobre el tema en cuestión.

Una buena forma de empezar suele ser estudiar aquellos factores geográficos que podrían indicar tendencias espaciales simples; por ejemplo, que los asentamientos se emplacen en torno a ciertos rangos altitudinales. Por ello, sería interesante definir toda una serie de condicionantes físicos y orográficos que busquen explicar las tendencias de localización generales. En segundo lugar, condicionantes que podemos denominar como potenciales, aquellos criterios que busquen asociar partes específicas del territorio o del paisaje con los sitios arqueológicos, como acumulaciones de agua natural o rutas de tránsito que discurren por los paisajes regionales.

Tanto los condicionantes físicos y orográficos como los potenciales pueden ser considerados como factores de primer orden, es decir, aquellos factores a través de los cuales podemos estudiar la distribución de los castros en el espacio. Desde aquí, podemos estudiar la intensidad (o densidad) de los asentamientos en la región seleccionada y dilucidar los efectos de los factores de primer orden, que, recordemos, se concretan en aquellos procesos o fenómenos externos a los asentamientos que provocan que su densidad varíe entre diferentes áreas. Finalmente, los efectos de primer orden pueden ser aislados para valorar los de segundo orden, o la interacción entre los propios asentamientos.

En este estudio sólo valoraremos la importancia de cuatro variables como efectos de primer orden, y se recogen en la **Tabla 5**.

Las diferentes covariables descritas en la **Tabla 5** pueden ser transformadas en mapas raster en un entorno SIG (por ejemplo, con *Quantum GIS*, *SAGA GIS*, o *GRASS GIS*, entre otros). Aunque este trabajo también puede realizarse enteramente en *R*, optaremos ahora por utilizar ambas opciones (*R* y *Quantum GIS*), con el objetivo de demostrar que muchas veces las soluciones más rápidas y sencillas se corresponden con un trabajo que integra el uso de ambas herramientas.

La altitud es una variable de sencilla creación a partir de un Modelo Digital del Terreno, por lo que no nos detendremos en su explicación. Por lo que respecta a la densidad de rutas de tránsito potencial, ya hemos descrito en el capítulo anterior cómo modelizar esta variable con *R Statistics* y la librería *leastcostpath* (Lewis 2020). Recordemos ahora que el método se centra en calcular la movilidad en todas las celdas de un MDT, generando rutas óptimas entre todos los pixeles que componen dicho MDT (**Figura 58: D**).

De igual forma que en el capítulo anterior, y dados los requerimientos computacionales de la aproximación descrita, procedemos a diseñar un método a partir de los límites del área de

	Covariables	Descripción
Factores de primer orden	Altitud (m.s.n.m.)	Modelo Digital de Elevaciones de 25m de resolución, construido a partir de datos LiDAR (sólo terreno)
	Densidad de rutas de tránsito potencial	Identificación de aquellas áreas del territorio que presentan mayor probabilidad estadística de ser transitadas, lo que las convierte en puntos clave o nodos teóricos en una red de tránsito peonil (Rodríguez Rellán, Fábregas Valcarce 2015)
	Prominencia topográfica	Definida como la función de la altura diferencial entre un individuo y el entorno tal y como es percibida desde el punto de vista del individuo en cuestión (Llobera 2001)
	Prominencia visual	Cálculo de la estructura visual inherente al paisaje (Llobera 2003, 2007; Llobera *et al.* 2010)

Tabla 5. Factores de primer orden definidos para este caso de estudio.

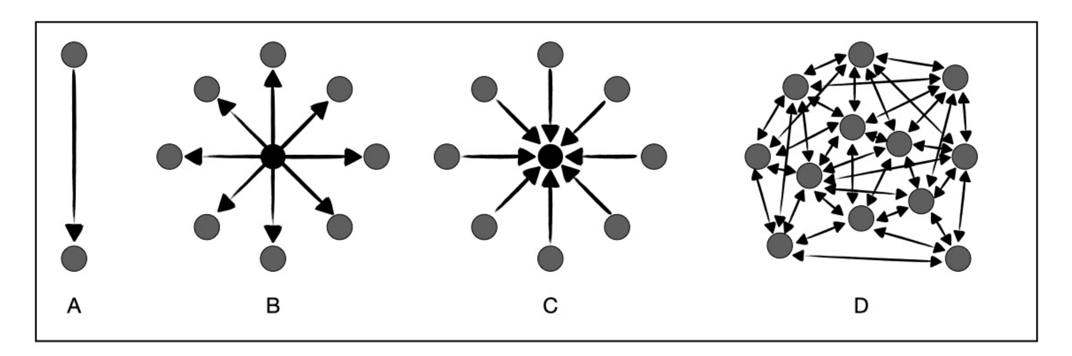


Figura 58. Opciones de conectividad para movimientos entre puntos. **A**: *one-to-one*; **B**: *one-to-many*; **C**: *many-to-one*; **D**: *many-to-many* (a partir de White, Barber 2012).

trabajo como puntos de origen y destino para las rutas óptimas (puntos regulares cada 100 m). Para evitar problemas de borde (véase al respecto Conolly, Lake 2009; Olaya 2014), las rutas óptimas se calcularán para un área más grande que la definida para el estudio de los efectos de primer y segundo orden.

En concreto, la aproximación que definimos se estructura en torno a las siguientes partes:

1. Creación del raster de coste, que identifica aquellas áreas del territorio por las que sería más fácil y difícil transitar, y aquellas por las que no se podría (por ejemplo, ríos de gran caudal, etc.). En aras de mantener el caso de estudio sencillo, aquí sólo utilizamos la pendiente y elevación como principales impedimentos del movimiento a través del paisaje.
2. Creación de las rutas óptimas desde los extremos del área de estudio (puntos de origen y destino situados de forma regular, cada 100 metros) (véase **Figura 59**).
3. Cálculo de la densidad de rutas óptimas (**Figura 60**).

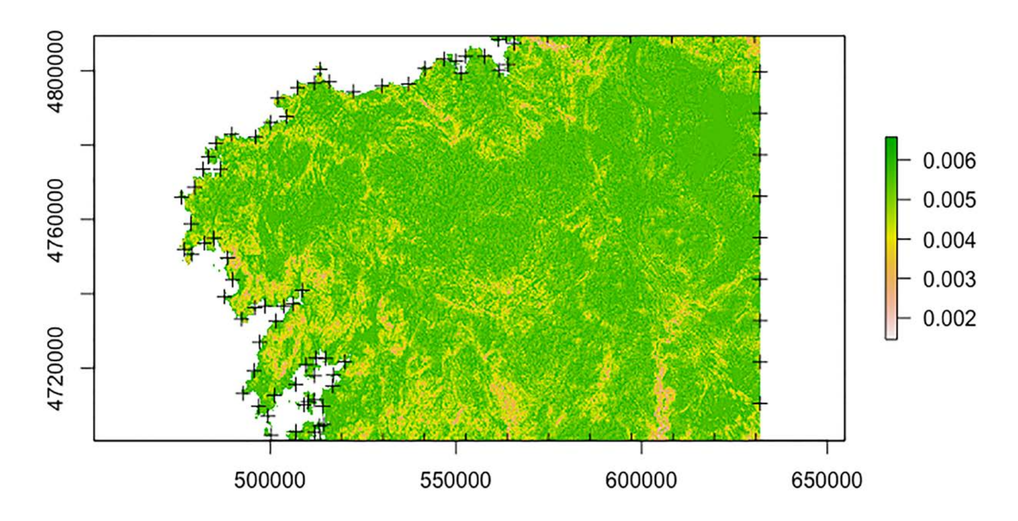


Figura 59. Mapa raster de coste del movimiento con los puntos de origen y destino para el cálculo de las rutas óptimas.

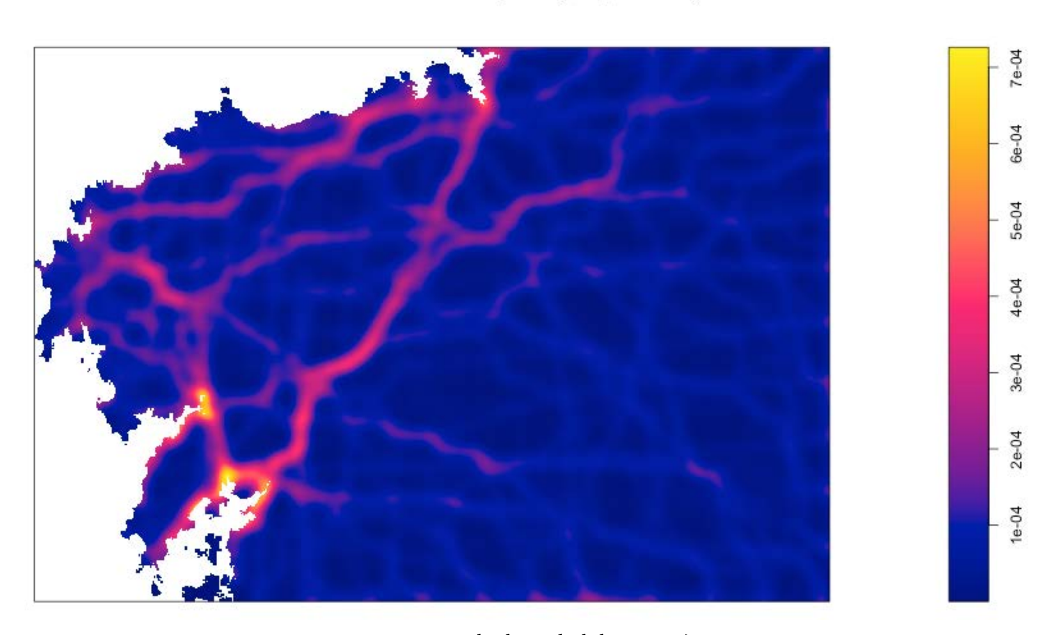


Figura 60. Mapa raster de densidad de rutas óptimas para el Noroeste de Galicia.

El cálculo de prominencia topográfica puede realizarse de forma sencilla en *SAGA GIS* a través del *plug-in Topographic Prominence Index* (o TPI). M. Llobera (2001) define este cálculo como la función de la altura diferencial entre un individuo y su entorno, tal y como es percibida desde el punto de vista del individuo en cuestión (**Figura 61**).

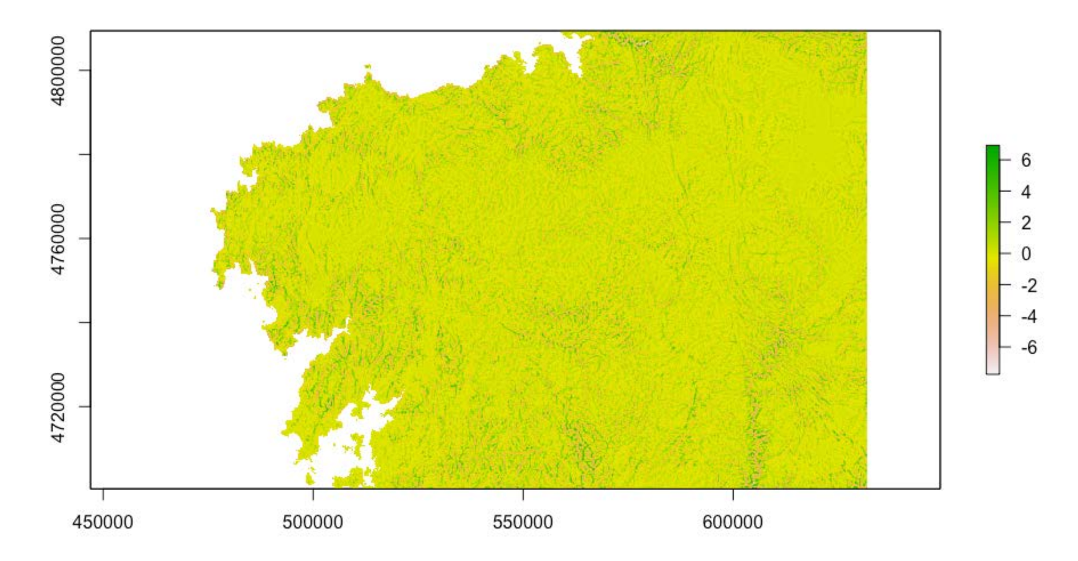


Figura 61. Mapa raster de la prominencia topográfica del terreno para el Noroeste gallego.

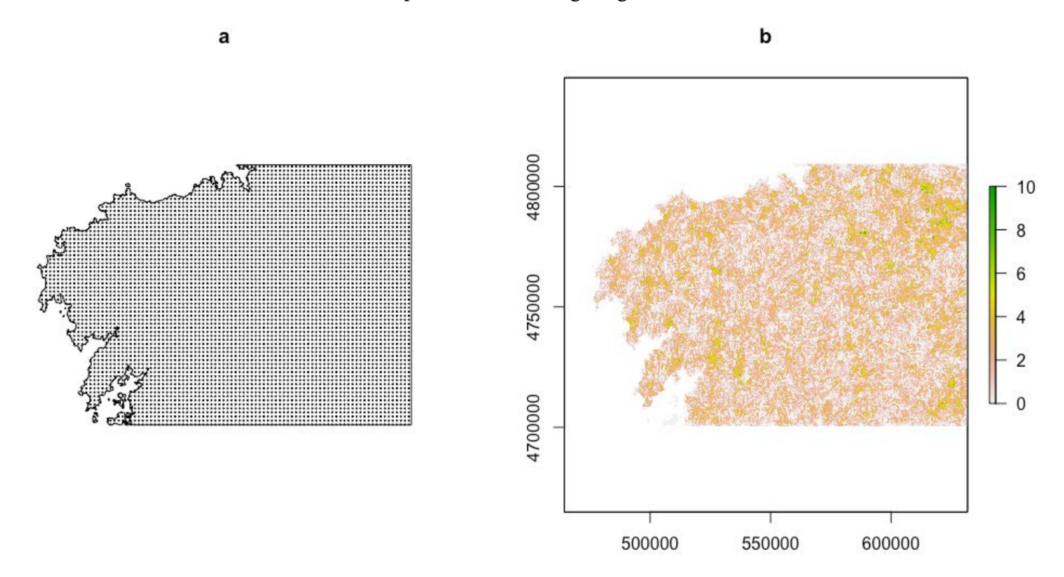


Figura 62. A: Malla regular de puntos de observador (cada 500m). **B:** Mapa raster de prominencia visual.

Por lo que respecta a la prominencia visual, esta variable se refiere al cálculo de la estructura visual inherente al paisaje, en los términos definidos también por M. Llobera en sus trabajos (por ejemplo, Llobera *et al.* 2010). El *plug-in Advanced viewshed* de *Quantum GIS* (Čučković 2016) es una de las herramientas más utilizadas. En nuestro caso hemos decidido calcular la visibilidad acumulativa sobre una malla de puntos regulares situados cada 500 metros en el territorio (**Figura 62A**). Ello supuso calcular la visibilidad individual para un total de 63.762 puntos de observador que, sumados, permiten obtener un raster de prominencia visual (**Figura 62B**).

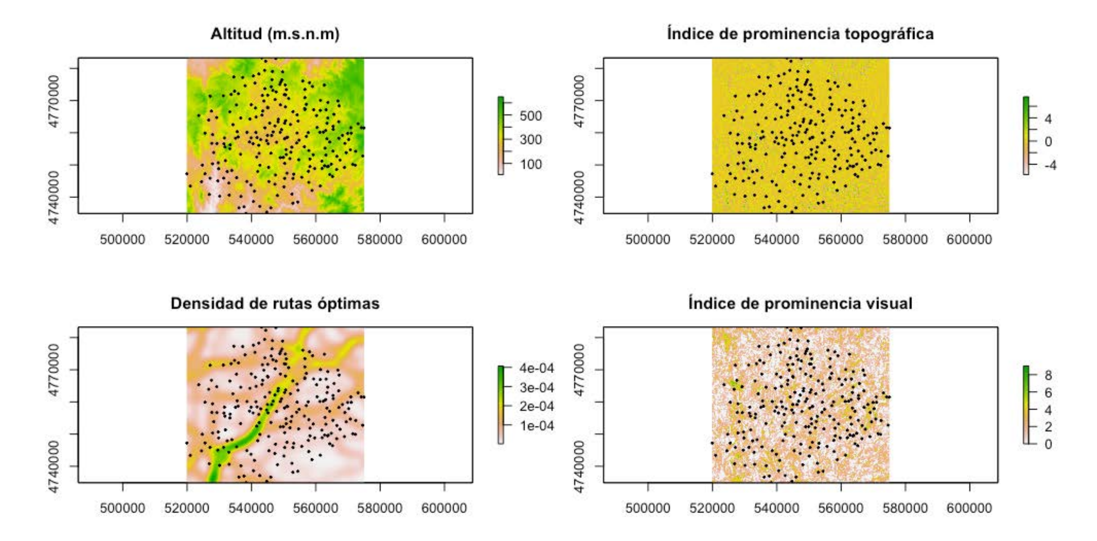


Figura 63. Covariables de primer orden seleccionadas para este caso de estudio. Los puntos negros se corresponden con los castros objeto de análisis.

Una vez hemos creado las diferentes covariables de primer orden con *R* y *Quantum GIS*, resumidas en la **Figura 63**, procederemos, a continuación, con el estudio de la distribución y los patrones monovariantes.

8.3. Estudio de la distribución de los castros: la intensidad espacial

El primer paso que suele realizarse es observar un mapa de distribución (**Figura 63**) para intentar concluir tendencias generales en la disposición espacial de los asentamientos. ¿Podría la distribución de castros ser el resultado de un proceso de aleatoriedad espacial completa? Aunque pueda parecer que dicha cuestión sea de Perogrullo, pues los procesos humanos que implican toma de decisiones nunca son aleatorios y suelen conllevar la consecución de efectos de primer, segundo y/o tercer orden, su análisis siempre es recomendable. Por poner un ejemplo irrelevante ahora, y bajo una lógica capitalista, pensemos en la localización de un *McDonalds* o una tienda de *Zara* en una gran ciudad. A nadie se le escapa que la localización de estos establecimientos dista mucho de ser aleatoria, pues suelen atender a dinámicas muy estudiadas de densidad de población, trasiego diario de gente, potencial económico del barrio, etc.

Uno de los métodos más utilizados para comprobar la hipótesis de aleatoriedad espacial completa en un conjunto de datos es la función K de Ripley. En el caso de los castros (**Figura 64**), el gráfico nos indica que estos yacimientos mantienen patrones diferentes según la escala, con valores en las distancias estudiadas (de 0 a 6000 m) que son mayores y menores en los castros (línea negra) que los esperados bajo condiciones aleatorias (línea roja punteada). En concreto, observamos una tendencia hacia la regularidad o dispersión entre distancias de ca. 800-1600 m, y agrupación a partir de los ca. 2000 m.

Sin embargo, la versión de la función K de Ripley que acabamos de utilizar en la figura anterior es la homogénea, es decir, el proceso de puntos aleatorio (Poisson) asume que la distribución

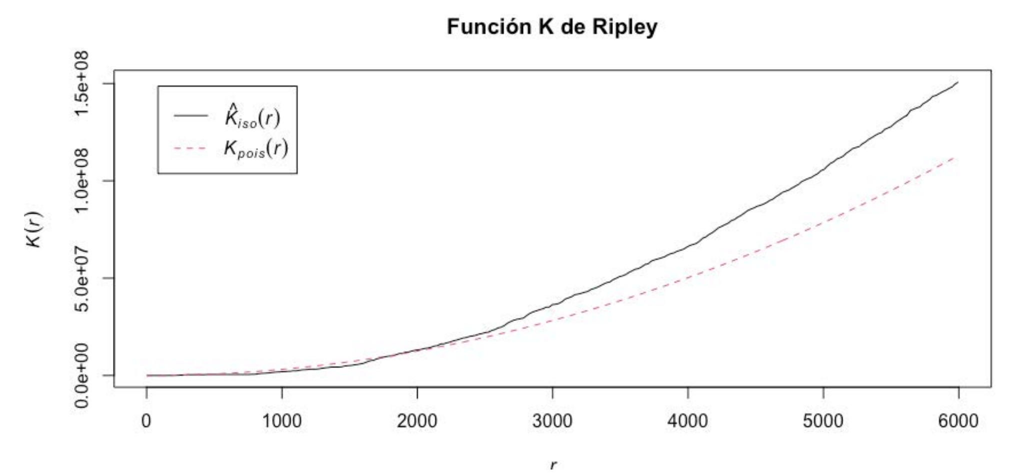


Figura 64. Función K de Ripley para los yacimientos castreños seleccionados.

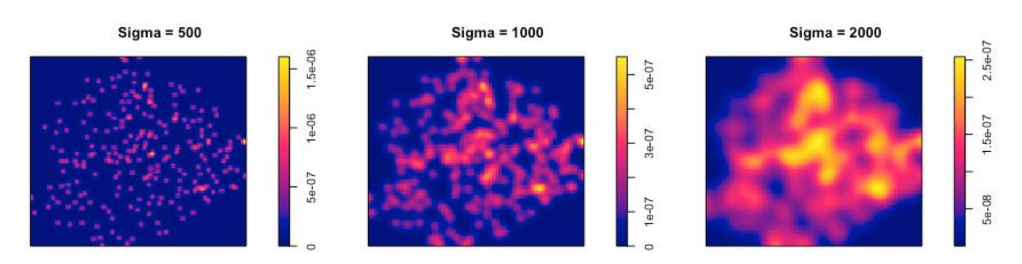


Figura 65. Análisis de densidad de los castros, con diferentes valores *sigma*.

de los castros no difiere en intensidad de un lugar a otro, lo cual sabemos que no es cierto. Los castros no se localizan en igual forma y distancia por el territorio, sino que suelen existir variaciones en su intensidad dependiendo de la zona de estudio. Por lo tanto, su distribución no es homogénea.

Para observar si la distribución de los castros no es homogénea podemos crear un mapa de densidad de puntos y evaluar visualmente el grado de concentración de castros y zonas de densidad elevada. Este análisis es útil como un primer paso, aunque los resultados dependen de un valor específico frecuentemente definido por el usuario, denominado *sigma*. El valor *sigma* define la escala de la interpolación utilizada para calcular la densidad, por lo que los resultados siempre cambian al modificarlo (**Figura 65**). Aunque hay formas de calcular automáticamente dicho valor (véase al respecto Baddeley *et al.* 2015: 171), es muy frecuente determinarlo manualmente y examinar visualmente diferentes resultados.

Los resultados permiten observar la existencia de tendencias no homogéneas, con mayor concentración de castros en el centro y norte del área de estudio que en los extremos. Un análisis que puede utilizarse para continuar estudiando la distribución es el test de cuadrantes (*quadrat test*). Tal y como indicamos en capítulos anteriores, esta prueba estadística se estructura en torno a la creación de una cuadrícula en el área de estudio con diferentes cuadrantes en su interior, y gracias a un test de chi-cuadrado se evalúa el grado en el que los

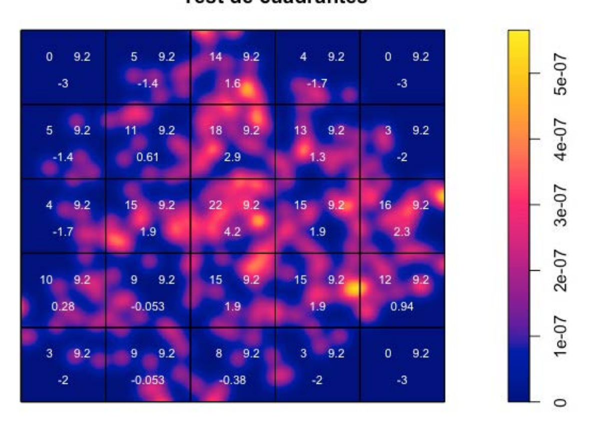


Figura 66. Test de cuadrantes basado en un cálculo de densidades (*sigma* = 1000).

Conditional Monte Carlo test of CSR using quadrat counts
Test statistic: Pearson X2 statistic
data: sppp_castros
X2 = 101.68, **p-value = 0.001**
alternative hypothesis: two.sided
Quadrats: 25 tiles (irregular windows)

Tabla 6. Prueba de chi cuadrado para el test de cuadrantes.

recuentos de los puntos difieren de los que cabría esperar bajo condiciones aleatorias (**Figura 66; Tabla 6**).

Ambos análisis sugieren que los cuadrantes observados (castros) difieren estadísticamente de los esperados bajo un proceso Poisson homogéneo (valor p = 0,001). Recordemos que el recuento observado (los castros que caen en cada cuadrícula) viene reflejado en el número de la parte superior izquierda, mientras que el esperado bajo condiciones *Poisson* es el de la parte superior derecha. En la parte inferior central de cada cuadrícula podemos encontrar los residuales, que nos indican cuándo los cuadrantes son relevantes (desviaciones mayores de -1 o 1). Como podemos observar, las áreas que poseen densidades elevadas de castros producen importantes desviaciones entre los datos observados y los esperados bajo condiciones aleatorias de homogeneidad. Esto nos indica que la distribución de los castros es poco homogénea en el área de estudio. Por lo tanto, lo más certero es pensar que existe algún tipo de condicionamiento externo al proceso de puntos observado (los castros) que provoca que su localización en el paisaje no sea homogénea (Bevan 2020: 63). Razón suficiente para prever la importancia de los efectos de primer y segundo orden, cuestión de la que nos ocuparemos en las líneas que siguen.

8.4. Patrones y efectos de primer orden: la importancia de las variables medioambientales

A la hora de estudiar los efectos de primer orden es recomendable estudiar el patrón monovariante de cada una de las variables para observar posibles tendencias en los datos. Se trataría de responder, por ejemplo, a la siguiente pregunta: ¿Están los castros localizados de

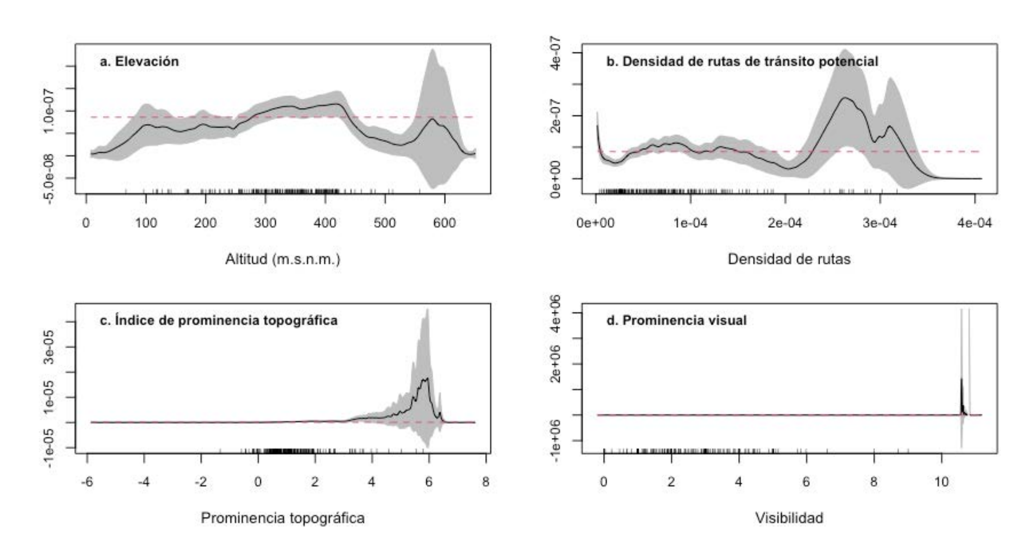


Figura 67. Gráficos de la función *rhohat* para las diferentes covariables.

forma preferencial en ciertos rangos altitudinales o visuales? Para responder a esta pregunta utilizaremos ahora la función *r.hohat* del paquete *spatstat* de *R*.[103] Esta función estima la relación entre la intensidad de un proceso puntual *Poisson* y una covariable espacial determinada (**Figura 67**).

La función *rhohat* es muy interesante porque nos permite observar tendencias monovariantes en los datos, anticipando posibles efectos de primer orden en la distribución de los yacimientos. En los gráficos de la **Figura 67** puede observarse la tendencia mostrada por los castros (las barras negras verticales, localizadas en la parte inferior) junto a un conjunto de simulaciones aleatorias esperadas, construidas bajo condiciones *Poisson* (la línea negra identifica el estadístico teórico esperado, mientras que el rango gris que lo envuelve se corresponde con un intervalo de confianza del 95%). Aparte de poder observar las tendencias generales de la localización de los yacimientos, nos interesan aquellas partes en las que el intervalo de confianza es más estrecho porque indican que la intensidad de los datos observados es mayor (Baddeley *et al.* 2015).

De la **Figura 67** podemos concluir que los castros suelen estar agrupados en torno a los 300-400 m.s.n.m., en áreas de reducida densidad de rutas de tránsito natural, pero ubicaciones caracterizadas por posicionamientos con prominencia topográfica y prominencia visual, aunque esta última no es muy elevada. De las cuatro covariables que manejamos, visualmente observamos que dos de ellas (elevación y densidad de rutas de tránsito) parecen no predecir en buena forma la distribución de los castros. Esto puede observarse al mapear la densidad espacial de los castros existente si las diferentes covariables, individualmente, fuesen las únicas impulsoras de dicha densidad (**Figura 68**).

Para evaluar los efectos de primer orden podemos optar por realizar un modelo de regresión logística, que se utiliza para predecir el resultado de una variable categórica (la intensidad o

[103] https://www.rdocumentation.org/packages/spatstat/versions/1.64-1/topics/rhohat

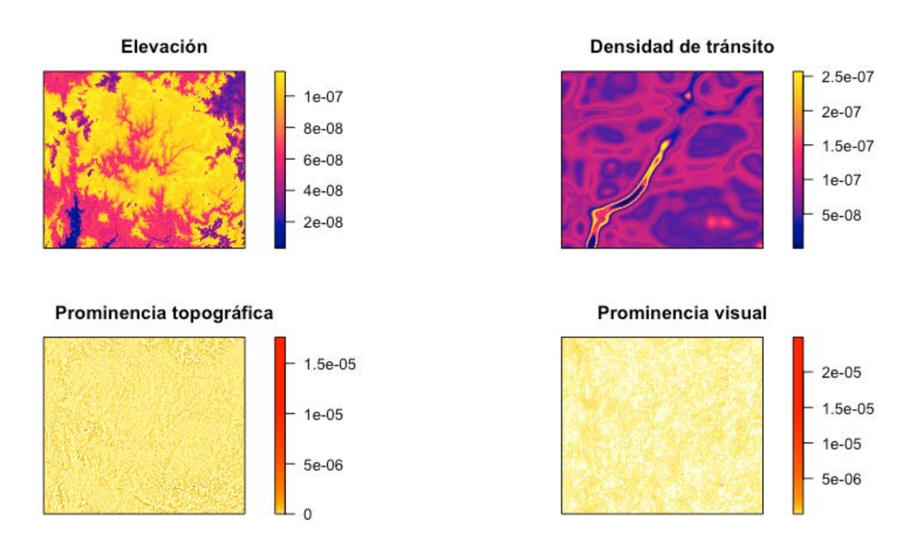


Figura 68. Densidad de castros predicha si las diferentes covariables fuesen las únicas impulsoras del proceso.

densidad de los castros) en función de un conjunto de variables independientes o predictoras (elevación, densidad de rutas óptimas, prominencia topográfica y prominencia visual). En nuestro caso, nos hemos basado en las investigaciones de C. Spencer y A. Bevan (2018), M. Carrero-Pazos, A. Bevan y M. Lake (2019) y P. Riris (2020) para realizar esta parte en *R Statistics*, y una buena descripción de estos métodos puede encontrarse en los trabajos de A. Bevan (*et al.* 2013; 2020).

Previo a la creación del modelo de regresión logística es conveniente realizar, tal y como señala P. Riris (2020), una prueba de colinealidad en las variables, por si fuese necesario eliminar alguna. Como los resultados de dicha prueba en nuestro caso son negativos (no existen variables colineales), procedemos a realizar la regresión logística con las cuatro covariables iniciales: elevación; densidad de rutas de tránsito potencial; prominencia topográfica; prominencia visual.

Las diferentes covariables serán ponderadas utilizando el criterio Akaike (AIC). Recordemos que, en términos generales, el AIC es una medida de la calidad relativa de un modelo estadístico para un conjunto de datos, y proporciona una estimación de la información que se pierde cuando se usa un determinado modelo para representar el proceso que ha generado los datos. En términos sencillos, el AIC es un método que permite seleccionar un modelo, y el mejor modelo será aquel que posea el AIC más bajo (Baddeley *et al.* 2015). Los resultados de la regresión logística minimizados con el criterio AIC se encuentran en la **Tabla 7**.

Los resultados indican que sólo la prominencia topográfica y la prominencia visual son las variables que predicen de forma significativa la distribución de los castros, eliminando la densidad de rutas y la elevación. Con ambas variables podemos construir la superficie de intensidad predictiva que contiene los efectos de primer orden (**Figura 69**).

El siguiente paso suele orientarse hacia investigar hasta qué punto la distribución espacial de los castros puede explicarse por el efecto de ambas covariables. Para ello, siguiendo a C.

Covariables	Estimación	Error estándar	Valor Z	Significativo
(Intercept)	-1.725045e+01	0.2873993473	-60.0225731	***
Prominencia topográfica	1.047142e+00	0.1027633476	10.1898379	***
Prominencia visual	2.861217e-01	0.0595812871	4.8022074	***
Elevación	-2.386203e-04	0.0007601891	-0.3138961	
*** p ≤ 0.0001 ** p ≤ 0.01 * p ≤ 0.05 (Sin símbolo) p > 0.05 Variable descartada: Densidad de tránsito				

Tabla 7. Resultados del modelo de regresión logística. Covariables ajustadas para el modelo de primer orden.

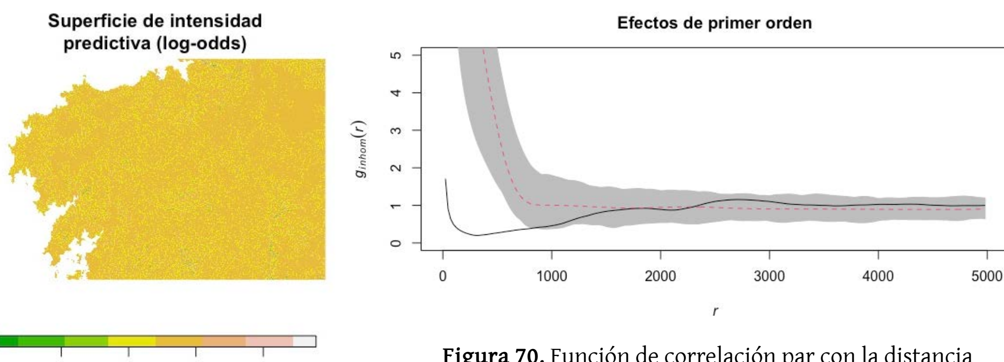


Figura 69. Superficie de intensidad de primer orden.

Figura 70. Función de correlación par con la distancia entre castros (línea negra) y distancias esperadas bajo los efectos de primer orden. El intervalo de confianza (rango gris) está creado con 10.000 simulaciones aleatorias de Monte Carlo, condicionadas por los efectos de primer orden.

Spencer y A. Bevan (2018) y A. Bevan (2020), podemos utilizar una función de correlación par en la que las simulaciones estadísticas que componen el rango aleatorio estén condicionadas por los efectos de primer orden (**Figura 70**).

Los resultados de la **Figura 70** muestran que los castros se encuentran distribuidos de forma regular por el territorio hasta los 2200m, distancia a partir de la cual su patrón parece tornarse agrupado. Y este patrón agrupado (ca. 2200 m, en adelante) parece explicarse por los efectos de primer orden. Resulta interesante destacar que la línea negra (castros) cae bajo el intervalo aleatorio sólo a partir de los 900 m, lo que nos indica que los efectos de primer orden pueden explicar sólo una parte de la tendencia espacial de la distribución de los castros. Es decir, a escalas locales (ca. 0-900 m), la regularidad observada no puede ser explicada por los efectos de primer orden, con lo que deben existir otros factores que están jugando un papel importante aquí. Quizás sean variables de primer orden que no se hayan tenido en cuenta en el análisis, o quizás dicha tendencia sea el resultado de una interacción específica entre los propios castros.

8.5. Patrones y efectos de segundo orden. Midiendo la territorialidad

Una vez estudiamos los efectos de primer orden y ponderamos su relevancia con respecto a la distribución de castros, estamos ahora en condiciones de estudiar los de segundo orden. Uno de los métodos más utilizados es ajustar modelos estadísticos teóricos a la distribución arqueológica, y observar hasta qué punto los datos observados encajan en dicho modelo o no (por ejemplo, Bevan *et al.* 2013, Bevan 2020). Este proceso, que en inglés se conoce como *fit models into data*, suele realizarse con simulaciones estadísticas conocidas que buscan estudiar la interacción entre los eventos de un proceso determinado (modelos de procesos de puntos tipo *Gibbs*, por ejemplo). En el caso de este trabajo existen dos modelos estadísticos de interacción de puntos pertenecientes a la familia *Gibbs* que nos permiten intentar explicar la regularidad a escalas locales observada en los castros:

- Modelo 1. Interacción de área (*Area Interaction model*) (Baddeley, Lieshout 1995). Este modelo crea una zona de influencia (o *buffer*) para los puntos de la distribución y genera múltiples patrones de inhibición y agrupación para los puntos dentro de esta área. En este caso, la interacción entre los puntos es mixta, entre repulsión y agrupación.
- Modelo 2. Proceso *hardcore* (*Hardcore model*) (Baddeley *et al.* 2015: 487). Este modelo genera una suerte de barrera o territorio alrededor de cada punto, zona en la cual ningún otro punto podrá caer.

Ambos modelos asumen interacción entre los puntos, bien de atracción o repulsión en el primer caso, bien de repulsión solamente en el segundo, sobre un radio o área determinada. Este es un aspecto muy interesante porque nos permite comprobar y cuantificar la existencia de una posible relación de territorialidad entre los yacimientos de nuestra área de estudio.

Los resultados de la aplicación de ambas simulaciones pueden observarse a través de las funciones de correlación par de la **Figura 71**. Los gráficos han sido generados, en este caso, con el paquete *onpoint* de *R*,[104] que incluye una barra inferior coloreada para identificar, visualmente, la tendencia del patrón. Incluimos también la función de correlación par con los efectos de primer orden para comparación.

Los resultados permiten observar que sólo cuando tenemos en cuenta los efectos de primer y segundo orden las tendencias de regularidad y agrupación observadas en el patrón espacial de los castros consiguen explicarse en su totalidad. Fijémonos en el caso de la **Figura 71B**, en la que la línea negra cae dentro del rango gris compuesto por 10.000 simulaciones aleatorias, simulaciones condicionadas tanto por los efectos de primer orden como por el modelo de interacción de área de segundo orden.

Si bien es cierto que en el caso del modelo de repulsión *hardcore* continúan existiendo distancias residuales que no consiguen explicarse por los efectos de primer orden ni por los de segundo (véase la parte amarilla del gráfico, en la **Figura 71C**), con elevada probabilidad esto significa que necesitamos prestar más atención a los efectos de primer orden. Convendría, pues, incluir más covariables o definir alguna que refleje mejor la intensidad de los castros en la zona.

[104] https://r-spatialecology.github.io/onpoint/

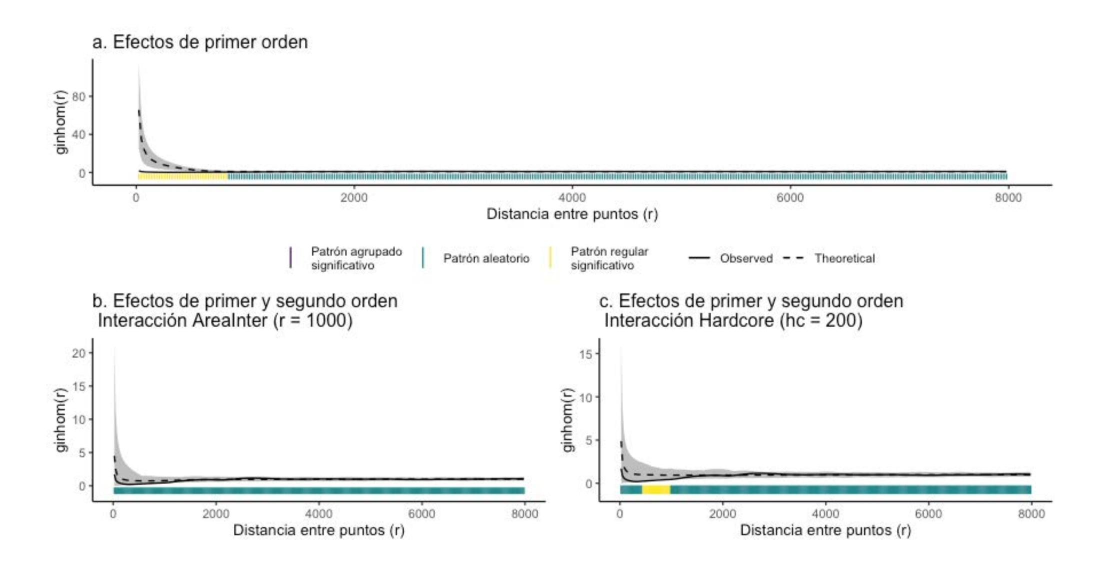


Figura 71. Funciones de correlación par con los diferentes modelos descritos anteriormente. **A:** Efectos de primer orden; **B:** Efectos de primer y segundo orden (modelo de interacción de área); **C:** Efectos de primer y segundo orden (modelo de interacción *hardcore*).

8.6. Conclusiones

El ejercicio realizado en este capítulo es sumamente útil a la hora de entender el patrón espacial del conjunto de yacimientos seleccionado. Teniendo en cuenta los resultados de la regresión logística realizada, la localización de los castros sobre zonas prominentes del paisaje y áreas de amplia visibilidad son variables que jugaron un papel un papel relevante en la conformación de los paisajes castreños de esta área de estudio. Quizás haya que entender esto a partir de conexiones visuales entre los propios castros, control de áreas específicas del territorio o, en definitiva, a partir de decisiones locacionales que parecen ser muy similares entre unas zonas y otras.

Por otra parte, el estudio de los efectos de segundo orden ha evidenciado la existencia de un patrón regular en la localización de los castros a escalas locales, y los diferentes modelos de simulación utilizados permiten definir y argumentar, en términos cuantitativos, que los castros analizados poseían una suerte de territorio de control a su alrededor, probablemente circunscrito a su entorno más cercano que a largas distancias (1 km según el modelo de interacción de área). Conseguimos, de esta forma, aportar argumentos cuantitativos a un posible modelo de territorial para un sector específico de los castros gallegos.

CUARTA PARTE:

REFLEXIONES FINALES

Capítulo 9.

Un futuro para la Arqueología computacional del territorio y paisaje: SIG de código libre y simulación estadística

La Arqueología computacional vive en la actualidad tiempos prósperos, algo visible en que cada vez son más los investigadores que incorporan a su formación técnicas informáticas, digitales o virtuales (Remondino, Campana 2014; Jiménez-Badillo 2017). Las temáticas se están diversificando, a la par que la introducción de nuevas herramientas permite manejar grandes cantidades de información, entrenar algoritmos matemáticos que detectan, automáticamente, estructuras arqueológicas en superficie, analizar digitalmente rocas con grabados que son invisibles al ojo humano o realizar simulaciones que abren el camino a nuevas teorías sobre el pasado más remoto. En este panorama, los modelos de simulación, estadísticos, de redes, de base agente y los relacionados con la inteligencia artificial, están llamados a convertirse en herramientas esenciales en las investigaciones arqueológicas del futuro más inmediato (Lake 2014; Rubio-Campillo 2017).

Como corriente de investigación, la Arqueología computacional no es una disciplina nueva, sino que cuenta con una arraigada tradición que puede retrotraerse a inicios de los años noventa en nuestro país, y que ha resultado en la formación más reciente de una nueva generación de arqueólogos especializados en técnicas computacionales.

Otro de los marcadores que nos indica que esta disciplina está en continua expansión es el éxito de congresos y reuniones especializadas, caso del clásico simposio *Computer Applications and Quantitative Methods in Archaeology* (CAA) que lleva celebrándose en diferentes ciudades desde 1973. Se trata, sin lugar a dudas, de la principal convención sobre métodos y técnicas computacionales en Arqueología.

Más recientemente asistimos también a la proliferación de revistas científicas especializadas, como el *Journal of Computer Applications in Archaeology*, *Virtual Archaeology Review*, o *Digital Applications in Archaeology and Cultural Heritage*. Todas ellas gozan de un volumen de publicación numeroso y sostenible, muestra de la salud de la que goza este campo de trabajo. También es muy común encontrarse con publicaciones que desarrollan métodos y técnicas computacionales en revistas más generales, como *Journal of Archaeological Science* o *Journal*

of Archaeological Method and Theory, por citar algunas de las más importantes en el campo arqueológico. Algunas de las revistas con mayor factor de impacto, tipo *Nature* o *Science* ya no aceptan trabajos que no incluyan metodologías reproducibles o en las que no se describan correctamente los métodos utilizados.

Vistas así las cosas, la Arqueología computacional del territorio y del paisaje, disciplina que, desde el punto de vista de este libro, se ocupa de estudiar la formación y evolución de los paisajes culturales pretéritos con métodos computacionales, tiene un futuro muy prometedor. Y es que las bases ya están sentadas. Contamos con más de cuatro décadas de desarrollo teórico y metodológico en tecnologías informáticas hoy democratizadas, como los Sistemas de Información Geográfica o los lenguajes de programación, que permiten a los arqueólogos manejar, de forma automática, información de diverso tipo y tamaño, generar cartografía analítica y beber de un sinfín de trabajos ya publicados sobre casuísticas específicas. Estamos en un momento idóneo para revisar hipótesis clásicas desde el punto de vista cuantitativo, depurar discursos tradicionales y proponer nuevos caminos.

Partiendo de esto, quienes que se dediquen a temas computacionales deben aprovechar los desarrollos informáticos de su tiempo y adaptar sus objetivos teóricos y metodológicos a las máquinas informáticas actuales. Hace diez años era muy difícil realizar un cálculo de visibilidad total. Hoy este trabajo suele ser asumible por un ordenador portátil al uso, por lo que los estudios de visibilidad de sitios arqueológicos deben considerar estas aproximaciones que, por definición, son más precisas que otros métodos.

En términos humanos, siguiendo a B. Marwick (2016: 445), desde el punto de vista tradicional los investigadores en la disciplina arqueológica e histórica poseen conocimientos profundos en ciertos dominios histórico-arqueológicos específicos, pero ahora es necesario que también desarrollen habilidades avanzadas en computación. Y es que, dice Nielsen (tomado de Rubio-Campillo 2017: 55), la interdisciplinariedad funciona realmente cuando hay un lenguaje común entre las disciplinas. El conjunto de la Ciencia ha encontrado en las matemáticas este lenguaje que permite y potencia el trabajo en común. En el caso de la Arqueología esto no ocurría, aunque en la actualidad el panorama está cambiando y cada vez es más frecuente encontrar perfiles de arqueólogos capaces de hablar este lenguaje común. Todo ello nos sitúa ante un evidente proceso de despegue y consolidación de la Arqueología computacional en nuestro país que indefectiblemente deberá reflejarse en los planes académicos universitarios, en los que materias como Arqueología cuantitativa, Arqueología computacional o Sistemas de Información Geográfica aplicados a Arqueología todavía siguen siendo *rara avis*. Sin lugar a dudas, y así lo hemos puesto de relieve en este monográfico, la simulación y el modelado estadístico, a través de lenguajes de programación como puede ser *R Statistics*, han llegado a la investigación arqueológica para quedarse.

Índice de Figuras

Índice de Tablas

Bibliografía

Ackermann, F. 1999. Airborne laser scanning-present status and future expectations. *Journal of Photogrammetry and Remote Sensing* 54: 64-67.

Aldalur, B., Santamaría, M. 2002. Realce de imágenes: filtrado especial. *Revista de Teledetección* 17: 31-42.

Aldenderfer, M., Maschner, H. D. G. 1996. *Anthropology, Space, and Geographic Information Systems*. Oxford: Oxford University Press.

Aldrich, J. H., Nelson, F. D. 1984. *Linear probability, logit and probit models*. Beverly Hills: Sage.

Allen, K. M. S., Green, S. W., Zubrow, E. B. 1990. *Interpreting space: GIS and archaeology*. London: Taylor & Francis.

Anderson, R. C. 1979. A kite-supported system for remote aerial photography. *Aerial Archaeology* 4: 4-7.

Anselin, L. 1995. The local indicators of spatial association: Lisa. *Geographical Analysis* 27: 93-115.

Anthony, D. 1990. Migration in Archeology: The Baby and the Bathwater. *American Anthropologist* 92(4): 895-914.

Ayán Vila, X., Amado Reino, X. 2001. La evaluación del impacto arqueológico de la red de gasificación de Galicia: consideraciones metodológicas y síntesis de resultados. *Trabajos de Prehistoria* 58: 127-141.

Baddeley, A. J., van Lieshout, M. N. M. 1995. Area-Interaction point processes. *Annals of the Institute of Statistical Mathematics* 47: 601–619.

Baddeley, A. J., Turner, R. 2005. Spatstat: An R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software* 12(6): 1-42.

Baddeley, A. J. 2010. *Analysing spatial point patterns in R*. Perth: CSIRO and University of Western Australia.

Baddeley, A. J., Rubak, E., Turner, R. 2015. *Spatial Point Patterns. Methodology and Applications with R*. Boca Raton: CRC Press.

Baena Preysler, J., Blasco Bosqued, M. C., Quesada Sanz, F. 1997. *Los SIG y el análisis espacial en Arqueología*. Madrid: UAM.

Bailey, T., Gatrell, T. 1995. *Interactive Spatial Data Analysis*. Harlow: Longman.

Baker, M. 2016. Reproducibility: Seek out stronger science. *Nature Communications* 537: 703–704.

Balbo, A. L., Rubio-Campillo, X., Rondelli, B., Ramírez, M., Lancelotti, C., Torrano, A., Salpeteur, M., Lipovetzky, N., Reyes-García, V., Montañola, C., Madella, M. 2014. Agent-based Simulation

of Holocene Monsoon Precipitation Patterns and Hunter-gatherer Population Dynamics in Semi-arid Environments. *Journal of Archaeological Method and Theory* 21: 426-446.

Bandera, M., Chaves Tristán, F. 1994. *Métodos analíticos y su aplicación a la arqueología*. Madrid: Editorial Gráficas Sol.

Barbeito Pose, V. J., Fábregas Valcarce, R., Rodríguez Rellán, C., Blanco Chao, R., Costa-Casais, M., Martín Seijo, M., Paz Camaño, A., Fariña Costa, A., Gorgoso López, L. 2015. Ocupacións domésticas na Serra do Barbanza: Resultados preliminares. *Gallaecia* 34: 125-158.

Barceló, J. A. 1991. *Arqueología, lógica y estadística*. Barcelona: Bellaterra.

Barceló, J. A. 1997. *Arqueología automática. Inteligencia artificial en arqueología*. Sabadell: Editorial Ausa.

Barceló, J. A., Briz, I., Vila, A. 1998. *New techniques for old times. Computer Applications in Archaeology*. Oxford: Archaeopress.

Barceló, J. A. 2007. *Arqueología y estadística 1. Introducción al estudio de la variabilidad de las evidencias arqueológicas*. Barcelona: Servei de Publicacions Universitat Autònoma de Barcelona.

Barceló, J. A. 2008. *Computational Intelligence in Archaeology*. New York: Information Science Reference.

Barceló, J. A., Bogdanovic, I. 2014. *Mathematics and Archaeology*. Boca Raton: CRC Press.

Barceló, J. A., Del Castillo Bernal, F. I. 2015. *Simulating Prehistoric and Ancient Worlds*. Berlin: Springer Verlag.

Baxter, M. 1994. *Exploratory Multivariate Analysis in Archaeology*. Edinburgh: Edinburgh University Press.

Baxter, M., Cool, H. 2016. *Basic Statistical Graphics for Archaeology with R: Life Beyond Excel*. Nottingham: Barbican Research Associates Nottingham.

Bello Diéguez, J. M., Vázquez Varela, J. M., Criado Boado, F. 1982. Aproximación a un modelo económico-social del megalitismo del noroeste peninsular. *Brigantium* 3: 33-39.

Bennett, R., Welham, K., Hill, R. A., Ford, A. 2011. Making the most of airborne remote sensing techniques for archaeological survey and interpretation, en D. C. Cowley (ed.) *Remote Sensing for Archaeological Heritage Management*: 99-107. Hungary: Archaeolingua.

Bennett, R., Welham, K., Hill, R. A., Ford, A. 2012. Using LiDAR as part of a multi-sensor approach to archaeological survey and interpretation, en R. Opitz y D. Cowley (eds) *Interpreting archaeological topography - airborne laser scanning, aerial photographs, and ground observation*: 1-12. Oxford: Oxbow Books.

Berganzo-Besga, I., Orengo, H. A., Lumbreras, F., Carrero-Pazos, M., Fonte, J., Vilas-Estévez, B. 2021. Hybrid MSRM-Based Deep Learning and Multitemporal Sentinel 2-Based Machine Learning Algorithm Detects Near 10k Archaeological Tumuli in North-Western Iberia. *Remote Sensing* 13: 4181.

Bevan, A., Conolly, J. 2004. GIS, Archaeological Survey, and Landscape Archaeology on the Island of Kythera, Greece. *Journal of Field Archaeology* 29: 123-138.

Bevan, A., Conolly, J. 2006. Multiscalar approaches to settlement pattern analysis, en G. Lock y B. Molyneaux (eds) *Confronting scale in archaeology: issues of theory and practice*: 217-234. Springer: New York.

Bevan, A., Conolly, J. 2009. Modelling spatial heterogeneity and nonstationarity in artifact-rich landscapes. *Journal of Archaeological Science* 36(4): 956-964.

Bevan, A. 2011. Computational models for understanding movement and territory, en V. Mayoral Herrera y S. Celestino Pérez (eds) *Tecnologías de Información Geográfica y Análisis Arqueológico Del Territorio. Actas Del V Simposio Internacional de Arqueología de Mérida*: 384-393. CSIC. Instituto de Arqueología.

Bevan, A., Lake, M. (eds.) 2013. *Computational Approaches to Archaeological Spaces*. Walnut Creek, California: Left Coast Press.

Bevan, A., Crema, E., Xiuzhen, L., Palmisano, A. 2013. Intensities, interactions and uncertainties: some new approaches to archaeological distributions, en A. Bevan y M. Lake (eds) *Computational approaches to archaeological landscapes*: 27-52. Walnut Creek, California: Left Coast Press.

Bevan, A. 2015. The data deluge. *Antiquity* 89(348): 1473-1484.

Bevan, A. 2020. Spatial point patterns and processes, en M. Gillings, P. Hacigüzeller y G. Lock (eds) *Archaeological Spatial Analysis. A Methodological Guide*: 60-76. London: Routledge.

Bewley, R. H., Crutchley, S. P., Shell, C. A. 2005. New light on an ancient landscape: LiDAR survey in the Stonehenge World Heritage Site. *Antiquity* 79: 636-647.

Bivand, R. S., Pebesma, E., Gómez-Rubio, V. 2013. *Applied spatial data analysis with R*. New York: Springer.

Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A., Selige, T. 2002. Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation, en E. Micheli, F. Nachtergaele y L. Montanarella (eds), *Soil Classification 2001*: 213-222. Luxembourg: European Soil Bureau.

Bourgeois, J., Marc, M. 2003. *Aerial Photography and Archaeology 2003: A Century of Information*. Gent: Academia Press.

Bove, F. J. 1981. Trend surface analysis and the Lowly Classic Maya collapse. *American Antiquity* 46: 93-112.

Bradley, R., Harding, J., Mathews, M. 1993. The siting of prehistoric rock art in Galloway, south-west Scotland. *Proceedings of the Prehistoric Society* 59: 269-283.

Brandolini, F., Carrer, F. 2020. Terra, Silva et Paludes. Assessing the Role of Alluvial Geomorphology for Late-Holocene Settlement Strategies (Po Plain - N Italy) Through Point Pattern Analysis. *Environmental Archaeology* 26(5): 511-525.

Brophy, K., Cowley, D. 2005. *From the air: Understanding Aerial Archaeology*. Stroud: The History Press.

Burillo Mozota, F. 2013. La prospección de superficie. En M. García-Díez y L. Zapata (eds) *Métodos y técnicas de análisis y estudio en Arqueología Prehistórica. De lo técnico a la reconstrucción de los grupos humanos*: 15-38. Bilbao: Servicio Editorial de la Universidad del País Vasco.

Burmeister, S. 2000. Archaeology and Migration. Approaches to an Archaeological Proof of Migration. *Current Anthropology* 41: 539-567.

Burrough, P. A., McDonnell, R. A. 1998. *Principles of Geographical Information Systems*. Oxford: Oxford University Press.

Bustelo Abuín, J., Rodríguez Rellán, C., Fábregas Valcarce, R., Barbeito Pose, V. 2017. Alén da Serra. O fenómeno tumular na Península do Barbanza a través dos SIX e a estatística espacial. *Gallaecia* 36: 53-72.

Canosa-Betés, J. 2016. Border surveillance: Testing the territorial control of the Andalusian defense network in center-south Iberia through GIS. *Journal of Archaeological Science: Reports* 9: 416-426.

Carlson, D. L. 2017. *Quantitative Methods in Archaeology Using R*. Cambridge: Cambridge University Press.

Carrera Ramírez, F. 2011. *El arte parietal en monumentos megalíticos del Noroeste Ibérico*. Oxford: BAR International Series 2190.

Carrero Pazos, M., Vilas Estévez, B., Romaní Fariña, E., Rodríguez Casal, A. A. 2014. La necrópolis del Monte de Santa Mariña revisitada: aportaciones del LiDAR aéreo para la cartografía megalítica de Galicia. *Gallaecia* 33: 39-57.

Carrero-Pazos, M., Vilas-Estévez, B. 2016. The possibilities of the aerial LiDAR for the detection of Galician megalithic mounds (NW of the Iberian Peninsula). The case of Monte de Santa Mariña, Lugo. En S. Campana, R. Scopigno, G. Carpentiero y M. Cirillo (eds), *Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology*: 901-908. Oxford, Archaeopress.

Carrero-Pazos, M. 2017. *El fenómeno tumular y megalítico en Galicia: aportaciones desde los Sistemas de Información Geográfica y la estadística espacial para el estudio de los patrones de localización.* Santiago de Compostela: Universidad de Santiago de Compostela. Tesis doctoral inédita. http://hdl.handle.net/10347/15823

Carrero-Pazos, M. 2018a. Modelando dinámicas de movilidad y visibilidad en los paisajes megalíticos gallegos. El caso del Monte de Santa Mariña y su entorno (Comarca de Sarria, Lugo). *Trabajos de Prehistoria* 75(2): 287-306.

Carrero-Pazos, M. 2018b. Beyond the scale. Building formal approaches for the study of spatial patterns in Galician moundscapes (NW Iberian Peninsula). *Journal of Archaeological Science: Reports* 19: 538-551.

Carrero-Pazos, M. 2023. Análisis de patrones espaciales de puntos para el estudio de tendencias locacionales en distribuciones de yacimientos arqueológicos. *Vegueta: Anuario de la Facultad de Geografía e Historia* 23(1): 97-114.

Carrero-Pazos, M., Bevan, A., Lake, M. 2019. The spatial structure of Galician megalithic landscapes (NW iberia): A case study from the Monte Penide region. *Journal of Archaeological Science* 108: 104968.

Carrero-Pazos, M., Bustelo-Abuín, J., Barbeito-Pose, V., Rodríguez-Rellán, C. 2020. Locational preferences and spatial arrangement in the barrow landscape of Serra do Barbanza (North-western Iberia). *Journal of Archaeological Science: Reports* 31: 102351.

Carrero-Pazos, M. 2021. La visibilidad como factor de localización en los megalitos del sur de Galicia. *Zephyrus* LXXXVII: 63-81.

Cerrato Casado, E. 2011. La prospección arqueológica superficial. Un método no destructivo para una ciencia que si lo es. *Arte, arqueología e historia* 18: 151-160.

Cerrillo-Cuenca, E. 2017. An approach to the automatic surveying of prehistoric barrows through LiDAR. *Quaternary International* 435: 135-145.

Cerrillo-Cuenca, E., Bueno-Ramírez, P. 2019. Counting with the invisible record? The role of LiDAR in the interpretation of megalithic landscapes in south-western Iberia (Extremadura, Alentejo and Beira Baixa). *Archaeological Prospection* 26: 251-264.

Cerrillo Cuenca, E., López López, A. 2020. Evaluación y perspectivas del uso del LiDAR en la arqueología española. *Boletín del Museo Arqueológico Nacional* 39: 221-238.

Chang, W. 2013. *R Graphics Cookbook*. Beijing: O'Reilly.

Chapman, H. 2006. *Landscape archaeology and GIS*. Stroud, Gloucestershire: Tempus.

Chapman, R. 1981. Archaeological theory and communal burial in Prehistoric Europe, en I. Hodder, G. Isaac y N. Hammond (eds) *Pattern to process. Studies in honour of David Clarke*: 387-412. Cambridge: Cambridge University Press.

Clark, P. J., Evans, F. C. 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology* 35(4): 445-453.

Clarke, D. L. 1977. *Spatial Archaeology*. London: Academic Press.

Comas, D., Ruiz, E. 1993. *Fundamentos de los Sistemas de Información Geográfica*. Barcelona: Ariel.

Conolly, J., Lake, M. 2006. *Geographic Information Systems in Archaeology*. Cambridge: Cambridge University Press.

Conolly, J., Lake, M. 2009. *Sistemas de Información Geográfica aplicados a la Arqueología*. Barcelona: Bellaterra.

Cordero Ruiz, T., Cerrillo Cuenca, E., Pereira, C. 2017. Detección de un nuevo campamento romano en las inmediaciones de Mérida mediante tecnología LiDAR. *SAGVNTVM. Papeles del Laboratorio de Arqueología de Valencia* 49: 197-201.

Costa García, J. M. 2015. Asentamientos militares romanos en el norte peninsular: aportes de la fotografía aérea histórica, la fotografía digital y el LiDAR aéreo. *Férvedes* 8: 25-44.

Costa-García, J. M. 2017. The potential of the Geographic Information Techniques for the analysis of the morphology and settlement patterns of the Roman military sites of early imperial era in Iberia, en V. Mayoral-Herrera, C. Parcero-Oubiña y P. Fábrega-Álvarez. *Archaeology and Geomatics. Harvesting the benefits of 10 years of training in the Iberian Peninsula (2006-2015)*: 209-226. Leiden: Sidestone Press.

Costa-García, J. M., Fonte, J. 2017. Scope and limitations of airborne LiDAR technology for the detection and analysis of Roman military sites in Northwest Iberia, en V. Mayoral-Herrera, C. Parcero-Oubiña y P. Fábrega-Álvarez. *Archaeology and Geomatics. Harvesting the benefits of 10 years of training in the Iberian Peninsula (2006-2015)*: 57-73. Leiden: Sidestone Press.

Cowley, D. 2012. In with the new, out with the old? Auto-extraction for remote sensing archaeology, en C. R. Bostater Jr., S. P. Mertikas, X. Neyt, C. Nichol, D. Cowley y J.-P. Bruyant (eds) *Proceedings Volume 8532, Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2012*: 853206. Edinburgh: International Society for Optics and Photonics.

Crema, E. R., Bevan, A., Lake, M. 2010. A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record. *Journal of Archaeological Science* 37: 1118-1130.

Crema, E. R. 2015. Time and Probabilistic Reasoning in Settlement Analysis, en J. A. Barceló e I. Bogdanovic (eds) *Mathematics and Archaeology*: 314-334. Boca Raton: CRC Press.

Cressie, N. 1991. *Statistics for Spatial Data*. Chichester: Wiley.

Criado Boado, F. 1984. El tercer factor o la lógica oculta del emplazamiento de los túmulos megalíticos gallegos. *Cuadernos de Estudios Gallegos* 35: 7-18.

Criado Boado, F., Aira Rodríguez, Mª J., Díaz-Fierros Viqueira, F. 1986. *La construcción del paisaje: Megalitismo y ecología. Sierra de Barbanza*. Santiago de Compostela: Xunta de Galicia, Consellería de Educación e Cultura, Dirección Xeral de Patrimonio Artístico e Monumental.

Criado Boado, F., Bonilla Rodríguez, A., Cerqueiro Landín, D., González Méndez, M., Méndez Fernández, F., Penedo Romero, R. 1988. Proyecto Bocelo-Furelos: arqueología del paisaje y prospección intensiva en Galicia. *Trabalhos de Antropologia e Etnologia* 28 (1-2): 241-250.

Criado Boado, F. 1988. *Contribución al estudio de las relaciones entre las comunidades megalíticas del noroeste peninsular y su medio natural: implicaciones socioeconómicas.* Santiago de Compostela: Universidade de Santiago de Compostela. Tesis doctoral inédita.

Criado Boado, F., Bonilla Rodríguez, A., Cerqueiro Landín, D., Díaz Vázquez, M., González Méndez, M., Infante Roura, F., Méndez Fernández, F., Penedo Romero, R., Rodríguez Puentes, E., Vaquero Lastres, J. 1991. *Arqueología del paisaje. El área Bocelo-Furelos entre los tiempos paleolíticos y medievales* (Vol. Arqueoloxía/Investigación: 6). A Coruña: Xunta de Galicia.

Criado Boado, F., Villoch Vázquez, V. 2000. Monumentalizing landscape: from present perception to the past meaning of Galician Megalithism (north-west Iberian Peninsula). *European Journal of Archaeology* 3(2): 188-216.

Criado Boado, F. 2017. A paisaxe como enxeñería inversa da materia e a memoria, en C. García García y M. García Miraz (eds) *Paisaxe e patrimonio: un percorrido polo territorio a través do Arquivo de Galicia*: 149-157. Santiago de Compostela: Consellería de Cultura, Educación e Ordenación Universitaria.

Crutchley, S., Crow, P. 2009. *Light Fantastic: Using airborne laser scanning in archaeological survey*. Swindon: English Heritage.
Čučković, Z. 2016. Advanced viewshed analysis: a Quantum GIS plug-in for the analysis of visual landscapes. *Journal of Open Source Software* 1(4)(32): 1-3.
Currás, B. X., Sastre, I. 2020. Egalitarianism and resistance: a theoretical proposal for Iron Age Northwestern Iberian archaeology. *Anthropol. Theory* 20: 300-329.
Dabezies Damboriarena, J. M. 2009. *La dimensión inmaterial del Paisaje. Una propuesta de documentación, caracterización y gestión del Patrimonio Cultural Inmaterial*. Porto Alegre: Pontíficia Universidade Católica do Río Grande do Sul.
Dacey, M. F. 1973. Statistical Tests of Spatial Association in the Locations of Tool Types. *American Antiquity* 38(3): 320-328.
De Reu, J., Bourgeois, J., De Smedt, P., Zwertvaegher, A., Antrop, M., Bats, M., Crombé, P. 2011. Measuring the relative topographic position of archaeological sites in the landscape, a case study on the Bronze Age barrows in northwest Belgium. *Journal of Archaeological Science* 38(12): 3435-3446.
Delibes de Castro, G. 1991. Megalitos ¿Todavía una civilización de muertos? *Arqrítica. Crítica de Arqueología Española* 2(2): 9-10.
Deuel, L. 1973. *Flights into Yesterday: the story of aerial archaeology*. Harmondsworth: Penguin.
Díaz Sanjurjo, M. 1904. Los caminos antiguos y el itinerario nº 18 de Antonino en la provincia de Orense. *Boletín de la Comisión de Monumentos Históricos y Artísticos de Orense* II: 37-43.
Diggle, P. 1983. *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.
Donnelly, K. 1978. Simulations to determine the variance and edge-effect of total nearest neighbour distance, en I. Hodder (ed) *Simulation methods in archaeology*. Cambridge: Cambridge University Press: 91-95.
Doran, J. E., Hodson, F. R. 1975. *Mathematics and Computing in Archaeology*. Harvard: Harvard University Press.
Drennan, R. D., Peterson, C. E. 2004. Comparing archaeological settlement systems with rank-size graphs: a measure of shape and statistical confidence. *Journal of Archaeological Science* 31: 533-549.
Drennan, R. D. 2009. *Statistics for Archaeology. A commonsense approach*. New York: Springer.
Ducke, B., Munch, U. 2005. Predictive modelling and the archaeological heritage of Brandenburg (Germany), en M. P. van Leusen, H. Kamermans y H. Amersfoot (eds), *Predictive Modelling for Archaeological Heritage Management: A Research Agenda*: 93-107. Amsterdam: National Service for Archaeological Heritage.
Duncan, R. B., Beckman, A. 2000. Site location in Pennsylvania and West Virginia, en K. L. Wescott, R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists. A Predictive Modeling Kit*: 33-58. London: Taylor & Francis.
Ebert, D. 2000. The state of the art in inductive predictive modeling: seven big mistakes (and lots of smaller ones), en K. L. Wescott, R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists. A Predictive Modeling Kit*: 129-134. London: Taylor & Francis.
Ebert, J. I., Kohler, A. 1988. The theoretical basis of archaeological predictive modeling and a consideration of appropriate data-collection methods, en W. J. Judge y L. Sebastian (eds) *Quantifying the Present and Predicting the Past. Theory, Method, and Application of Archaeological Predictive Modeling*: 97-171. Denver: US Department of the Interior.
Eguileta Franco, J. M. (1999). *A Baixa Limia na Prehistoria Recente. Arqueoloxía dunha Paisaxe na Galicia interior*. Ourense: Excma. Diputación Provincial de Ourense.
Epstein, J. M. 2008. Why Model? *Journal of Artificial Societies and Social Simulation* 11(4): https://www.jasss.org/11/4/12.html

Eve, S. J., Crema, E. R. 2013. A house with a view? Multi-model inference, visibility fields, and point process analysis of a Bronze Age settlement on Leskernick Hill (Cornwall, UK). *Journal of Archaeological Science* 43: 267-277.

Fábrega-Álvarez, P., Parcero Oubiña, C. 2007. Proposal for an archaeological analysis of pathways and movement. *Archeologia e Calcolatori* 18: 121-140.

Fábrega-Álvarez, P. 2017. *Recorriendo y observando paisajes digitales. Una aproximación al análisis arqueológico con Tecnologías de la Información Geográfica (TIG)*. Jaén: Universidad de Jaén. Tesis doctoral inédita.

Fábrega-Álvarez, P., Parcero Oubiña, C. 2019. Now you see me. An assessment of the visual recognition and control of individuals in archaeological landscapes. *Journal of Archaeological Science* 104: 56-74.

Fábregas Valcarce, R., Vilaseco Vázquez, X. I. 2015. Building forever or just for the time? en L. Laporte y C. Scarre (eds) *The Megalithic Architectures of Europe*: 101-110. Oxford: Oxbow Books.

Fairén, S. 2004. ¿Se hace camino al andar? Influencia de las variables medioambientales y culturales en el cálculo de caminos óptimos mediante SIG. *Trabajos de Prehistoria* 61: 25-40.

Fels, J. E. 1992. Viewshed simulation and analysis: an interactive approach. *GIS World Special Issue July*: 54-59.

Fernández Cacho, S. 2009. Bases conceptuales y metodológicas de los modelos predictivos en Arqueología. *E-Ph Cuadernos* 1: 8-32.

Fernández García, F. 2000. *Introducción a la fotointerpretación*. Barcelona: Ariel.

Fernández Martínez, V. 1989. La prospección arqueológica: planteamientos, ayudas y técnicas, en V. Fernández Martínez (ed.) *Teoría y método de la Arqueología*: 54-67. Madrid: Síntesis.

Ferreira de Sousa, M. G. 2013. *O fenómeno tumular e megalítico na região galaico-portuguesa do Minho*. Santiago de Compostela. Santiago de Compostela: Universidade de Santiago de Compostela. Tesis doctoral inédita.

Fischer, P. F. 1993. Algorithm and implementation uncertainty in viewshed analysis. *International Journal of Geographical Information Science* 7: 331-347.

Fischer, P. F., Farrelly, C., Maddocks, A., Ruggles, C. L. N. 1997. Spatial Analysis of visible areas from the Bronze Age cairns of Mull. *Journal of Archaeological Science* 24: 581-592.

Flannery, K. V., Coe, M. D. 1969. Social and Economic Systems in Formative Mesoamerica, en S. R. Binford y L. R. Binford (eds) *New Perspectives in Archaeology*: 267-284. Chicago: Aldine Transaction.

Fleming, A. 1973. Tombs for the living. *Man* 8(2): 177-193.

Fletcher, M., Lock, G. 1991. *Digging Numbers. Elementary Statistics for Archaeologists*. Oxford: Oxford University Press.

Fonte, J. 2015. *Paisagens em mudança na transição entre a Idade do Ferro e a Época Romana no Alto Tamega e Cávado*. Santiago de Compostela: Universidade de Santiago de Compostela. Tesis doctoral inédita.

Forte, M. 1995. Scientific visualization and archaeological landscape: The case study of a terramara, Italy, en G. Lock y Z. Stančič (eds), *Archaeology and Geographic Information Systems: A European Perspective*: 231-238. London: Taylor & Francis.

Fotheringham, S., Brundson, C., Charlton, M. 1999. *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage.

Fovet, É., Zakšek, K. 2014. Path network modelling and network of aggregated settlements: a case study in Languedoc (Southeastern France), en S. Polla y P. Verhagen (eds) *Computational approaches to the study of movement in archaeology. Theory, practice and interpretation of factors and effects of long-term landscape formation and transformation*: 43-72. Berlin: De Gruyter.

Fraser, D. 1988. The orientation of visibility from the chambered cairns of Eday, Orkney, en C. L. N. Ruggles (ed) *Records in Stone*: 325-337. Cambridge: Cambridge University Press.

Gaffney, V., van Leusen, M. 1995. Postscript-GIS, environmental determinism and archaeology: a parallel text, en G. Lock y Z. Stančič (eds) *Archaeology and Geographical Information Systems: A European Perspective*: 367-382. London: Taylor & Francis.

Gaffney, V., Stančič, Z., Watson, H. 1996. Moving from catchments to cognition: tentative steps toward a larger archaeological context for GIS, en M. Aldenderfer y H. D. G. Maschner (eds) *Anthropology, Space and Geographic Information Systems*: 132-154. Oxford: Oxford University Press.

García García, M. 2015. *Poblamiento y sociedad en el valle medio del Duero durante la Edad de los Metales (ca. 2800 - 450 cal. ANE). Un estudio de Arqueología del Paisaje.* Valladolid: Universidad de Valladolid. Tesis doctoral inédita.

García Sanjuán, L. 2004. La prospección arqueológica de superficie y los SIG, en J. C. Martín de la Cruz (ed.) *I Encuentro Internacional de Informática Aplicada a la Investigación y la Gestión Arqueológicas*: 185-209. Córdoba: Universidad de Córdoba.

García Sanjuán, L. 2005. *Introducción al Reconocimiento y Análisis Arqueológico del Territorio.* Barcelona: Ariel.

GDAL Development Team (2014). LiDAR Derived Digital Terrain Models. *Remote Sensing.*

Gearey, B. R., Chapman, H. 2006. Digital Gardening: an approach to simulating elements of palaeovegetation and some implications for the interpretation of prehistoric sites and landscapes, en T. L. Evans y P. Daly (eds) *Digital Archaeology: bridging method and theory*: 171-190. London: Routledge.

Geary, R. 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician* 5: 115-145.

Getis, A., Ord, J. K. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24: 189-206.

Gillings, M. 2009. Visual affordance, landscape and the megaliths of Alderney. *Oxford Journal of Archaeology* 28: 335-356.

Gillings, M. 2015. Mapping invisibility: GIS approaches to the analysis of hiding and seclusion. *Journal of Archaeological Science* 62: 1-14.

Gillings, M., Hacigüzeller, P., Lock, G. 2020. *Archaeological Spatial Analysis. A methodological guide*. London: Routledge.

González Insua, F. 2013. Aproximación a la distribución espacial de sitios arqueológicos de la Prehistoria Reciente en la Ría de Arousa (Pontevedra). *Gallaecia* 32: 129-167.

González-Ruibal, A. 2011. The politics of identity: ethnicity and the economy of power in Iron Age Northwest Iberia, en G. Cifani (ed.) *Landscape, Ethnicity, Identity in the Archaic Mediterranean Area*: 245-266. Oxford: Oxbow Books.

Goreaud, F., Pélissier, R. 2000. Spatial Structure Analysis of Heterogeneous Point Patterns: Examples of application to forest stands. *ADS in ADE-4 Topic Documentation* 8.1: 1-49.

Grau Mira, I. 2006. *La aplicación de los SIG en Arqueología del Paisaje*. Alicante: Publicaciones de la Universidad de Alicante.

Graves, D. 2011. The use of predictive modelling to target Neolithic settlement and occupation activity in mainland Scotland. *Journal of Archaeological Science* 38: 633-656.

Guisan, A., Weiss, S. B., Weiss, A. D. 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143: 107-122.

Hacigüzeller, P. 2012. GIS, critique, representation and beyond. *Journal of Social Archaeology* 12: 245-263.

Haining, R. 2003. *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University Press.

Hanks, B. 2008. The past in later Prehistory, en A. Jones (ed.) *Prehistoric Europe, Theory and Practice*: 255-284. Oxford: Blackwell.

Herzog, I., Posluschny, A. 2011. Tilt-slope-dependent least cost path calculations revisited, en E. Jerem, F. Redo y F. Szeverényi (eds) *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference Budapest, April 2-6, 2008*: 212-218. Budapest, Archaeolingua.

Herzog, I. 2013. The Potential and Limits of Optimal Path Analysis, en A. Bevan y M. Lake (eds) *Computational Approaches to Archaeological Spaces*: 179-211. Walnut Creek, California, Left Coast Press: 179-211.

Hesse, R. 2010. LiDAR-derived Local Relief Models – a new tool for archaeological prospection. *Archaeological Prospection* 17: 67-72.

Hesse, R. 2013. *Lidar Visualization Toolbox. Visual Basic.NET (version 1.0.0.20)*: http://sourceforge.net/projects/livt/?source=navbar

Hesse, R. 2016. Visualisierung hochauflösender Digitaler Geländemodelle mit LiVT, en U. Lieberwirth e I. Herzog (eds) *3D-Anwendungen in der Archäologie. Computeranwendungen und quantitative Methoden in der Archäologie. Workshop der AG CAA und des Exzellenzclusters Topoi 2013*. Berlin: Topoi.

Heywood, I. 1998. *Introduction to Geographical Information Systems*. Boston: Adisson Wesley Longman.

Hiller, J. K., Smith, M. 2008. Residual relief separation: digital elevation model enhancement for geomorphological mapping. *Earth Surface Processes and Landforms* 33: 2266–2276.

Hodder, I., Hassall, M. 1971. The Non-Random Spacing in Romano-British Walled Towns. *Man* 6: 391-407.

Hodder, I. 1972. Locational models and the study of romano-british settlement, en D. L. Clarke (ed.) *Models in Archaeology*: 887-909. London: Methuen.

Hodder, I., Orton, C. 1976. *Spatial Analysis in Archaeology*. Cambridge: Cambridge University Press.

Horn, B. K. P. 1981. Hill shading and the reflectance map. *Proceedings IEEE* 69(1): 14-47.

Humme, A., Lindenbergh, R., Sueur, C. 2006. Revealing Celtic fields from lidar data using kriging based filtering. *Proceedings of the ISPRS Commission V Symposium (Dresden, 25–27 September)* XXXVI: 1-5.

Ihm, P. (1978). *Statistik in der Archäeologie*. Bonn.

Izraelevitz, P. 2003. A fast algorithm for approximate viewshed computation. *Photogrammetric Engineering and Remote Sensing* 69: 767-774.

Jiménez-Badillo, D. 2017. *Arqueología computacional. Nuevos enfoques para la documentación, análisis y difusión del patrimonio cultural*. México: Instituto Nacional de Antropología e Historia.

Jorge, V. O. 1982. *Megalitismo do norte de Portugal: o distrito do Porto. Os monumentos e a sua problemática no contexto europeu*. Porto: Universidade do Porto. Tesis doctoral inédita.

Judge, W. J., Sebastian, L. 1988. *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Preditive Modeling*. Washington: US Bureau of Land.

Kamermans, H., Gojda, M., Posluschny, A. G. 2014. *A Sense of the Past. Studies in current archaeological applications of remote sensing and non-invasive prospection methods*. Oxford: BAR International Series.

Kelly, R. L. 1992. Mobility/sedentism: concepts, archaeological measures, and effects. *American Review of Anthropology* 21: 43-66.

Knitter D., Nakoinz, O. 2018. Point Pattern Analysis as Tool for Digital Geoarchaeology: A Case Study of Megalithic Graves in Schleswig-Holstein, Germany, en C. Siart, M. Forbriger y O. Bubenzer (eds) *Digital Geoarchaeology: New Techniques for Interdisciplinary Human-Environmental Research*. Cham: Springer.

Kokalj, Ž., Zakšek, K., Oštir, K. 2011. Application of sky-view factor for the visualisation of historic landscape features in lidar-derived relief models. *Antiquity* 85(327): 263-273.

Kokalj, Ž., Hesse, R. 2017. *Airborne laser scanning raster data visualization: A Guide to Good Practice*. Ljubljana: Založba ZRC.

Kristiansen, K. 2015. The decline of the Neolithic and the rise of Bronze Age Society, en C. Fowler, J. Harding y D. Hofmann (eds) *The Oxford Handbook of Neolithic Europe*: 1-19. Oxford: Oxford University Press.

Kvamme, K. 1983. Computer processing techniques for regional modeling of archaeological site locations. *Advances in Computer Archaeology* 1: 26-52.

Kvamme, K. 1990. The fundamental principles and practice of predictive archaeological modeling, en A. Voorrips (ed.) *Mathematics and Information Science in Archaeology: A Flexible Framework*: 257-295. Bonn: Holos-Verlag.

Kvamme, K. 2006. There and Back Again: Revisiting Archaeological Location Modeling, en M. W. Mehrer y K. L. Wescott (eds) *GIS and Archaeological Site Location Modeling*: 3-38. London: Taylor & Francis.

Lagerås, K. E. 2002. Visible intentions? Viewshed analysis of Bronze Age burial mounds in Western Scania, Sweden, en C. Scarre (ed.) *Monuments and landscape in Atlantic Europe, perception and society during the Neolithic and the Early Bronze Age*: 179-191. London: Routledge.

Lake, M., Woodman, P., Mithen, S. J. 1998. Tailoring GIS software for archaeological applications: an example concerning viewshed analysis. *Journal of Archaeological Science* 25: 27-38.

Lake, M., Woodman, P. 2003. Visibility studies in archaeology: a review and case study. *Environment and Planning B: Planning and Design* 30: 689-707.

Lake, M. 2012. Open archaeology. *World Archaeology* 44(4): 471-478.

Lake, M., Ortega, D. 2013. Compute-intensive GIS visibility analysis of the settings of prehistoric stone circles, en A. Bevan y M. Lake (eds) *Computational Approaches to Archaeological Spaces*: 213-242. Walnut Creek, California: Left Coast Press.

Lake, M. 2014. Trends in Archaeological Simulation. *Journal of Archaeological Method and Theory* 21: 258-287.

Landis, H. 2002. Production-ready global illumination. *Course 16 notes, SIGGRAPH. 2002*.

Last, J. 2007. *Beyond the grave, new perspectives on barrows*. Oxford: Oxbow Books.

Lee, J., Stucky, D. 1998. On applying viewshed analysis for determining least-cost paths on Digital Elevation Models. *International Journal of Geographical Information Systems* 12: 891-905.

Lewis, J. 2020. *Leastcostpath: Modelling Pathways and Movement Potential Within a Landscape* (version 1.8.1). https://cran.r-project.org/web/packages/leastcostpath/index.html

Llobera, M. 1996. Exploring the topography of mind: GIS, social space and archaeology. *Antiquity* 70: 612-622.

Llobera, M. 2000. Understanding movement: a Pilot Model Towards the Sociology of Movement, en G. Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies*: 66-84. Washington: IOS Press.

Llobera, M. 2001. Building past landscape perception with GIS: understanding topographic prominence. *Journal of Archaeological Science* 28: 1005-1014.

Llobera, M. 2003. Extending GIS-based visual analysis: The concept of "visualscapes". *International Journal of Geographical Information Science* 17: 25-48.

Llobera, M. 2006. Arqueología del Paisaje en el siglo XXI. Reflexiones sobre el uso de los SIG y modelos matemáticos, en I. Grau Mira (ed.) *La aplicación de los SIG en la arqueología del paisaje*: 109-124. Alicante: Servicio de Publicaciones Universidad de Alicante.

Llobera, M. 2007. Modeling visibility through vegetation. *International Journal of Geographical Information Science* 21: 799-810.

Llobera, M., Sluckin, T. J. 2007. Zigzagging: Theoretical Insights on Climbing Strategies. *Journal of Theoretical Biology* 249(7): 206-217.

Llobera, M., Wheatley, D., Steele, J., Cox, S., Parchment, O. 2010. Calculating the Inherent Visual Structure of a Landscape ("Total Viewshed") Using High-Throughput Computing, en F. Niccolucci y S. Hermon (eds) *Beyond the Artefact. Digital Interpretation of the Past. Proceedings of CAA2004 (Prato, 13-17 April 2004)*: 146-151. Budapest: Archaeolingua.

Llobera, M. 2011. Archaeological Visualization: Towards an Archaeological Information Science (AISc). *Journal of Archaeological Method and Theory* 18(3): 193-223.

Llobera, M. 2015. Working the Digital: Some thoughts from Landscape Archaeology, en R. Chapman y A. Wylie (eds) *Material Evidence: Learning from Archaeological Practice*: 173-188. Abingdon: Routledge.

Lock, G., Stančič, Z. 1995. *Archaeology and Geographical Information Systems: An European Perspective*. London: CRC Press.

Lockyear, K., Rahtz, S. 1991. *Computer applications and quantitative methods in archaeology*. Oxford: BAR International Series.

López Romero, E., Mañana Borrazás, P. 2013. *El círculo lítico de Monte Lobeira (Vilanova de Arousa, Pontevedra). Trabajos de 2008 y 2010*. Santiago de Compostela: Incipit, CSIC.

López-Felpeto Gómez, M. A. 2014. *Localización e catalogación do patrimonio arqueolóxico no ámbito do monte veciñal e man común de Santo Estevo de Parga (Guitiriz, Lugo).* Santiago de Compostela: Dirección Xeral de Patrimonio Cultural, Xunta de Galicia. Informe valorativo inédito.

Maciñeira, F. 1935. La distribución de las estaciones prehistóricas ortegaleses y sus características. *Boletín de la Real academia Galega* 259: 169-181.

Makarowicz, P., Goslar, T., Niebieszczański, J., Cwaliński, M., Kochkin, I. T., Romaniszyn, J., Lysenko, S. D., Ważny, T. 2018. Middle Bronze Age societies and barrow line chronology. A case study from the Bukivna 'necropolis', Upper Dniester Basin, Ukraine. *Journal of Archaeological Science* 95: 40-51.

Manly, B. F. J. 1991. *Randomization and Monte Carlo Methods in Biology*. London: Chapman and Hall.

Mañana Borrazás, P., Gianotti García, C., González Insua, F., Caramés Moreira, V. (2010). Aplicación de tecnologías geoespaciales para la documentación del círculo lítico de Monte Lobeira, Vilanova de Arousa (Pontevedra). *Cuadernos de Estudios Gallegos* 57(123): 25-52.

Márquez Pérez, J., Vallejo Villalta, I., Álvarez Francoso, J. I. (2015). Estimación del tiempo de demora en rutas pedestres: comparación de algoritmos. *GeoFocus* 15: 47-74.

Martín de la Cruz, J. C., Lucena Martín, A. M.ª 2004. *I Encuentro Internacional de Informática Aplicada a la Investigación y la Gestión Arqueológica*. Córdoba: Servicio de Publicaciones de la Universidad de Córdoba.

Martínez de Padín, L. 1849. *Historia política, religiosa y descriptiva de Galicia*. Madrid: Establecimiento tipográfico de A. Vicente.

Martínez del Pozo, J. A., Cerrillo Cuenca, E., Salas Tovar, E. 2013. Low altitude aerial photography applications for digital surface models creation in archaeology. *Transactions in GIS* 18(2): 227-246.

Martinón-Torres, M. 2001. Los megalitos de término. Crónica del valor territorial de los monumentos megalíticos a partir de las fuentes escritas. *Trabajos de Prehistoria* 58(1): 95-108.

Marwick, B. 2016. Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory* 24: 424–450.

Marwick, B., d'Alpoim Guedes, J., Barton, C. M., Bates, L. A., Baxter, M., Bevan, A., Bollwerk, E. A., Bocinsky, R. K., Brughmans, T., Carter, A. K., Conrad, C., Contreras, D. A., Costa, S., Crema, E. R., Daggett, A., Davies, B., Drake, B. L., Dye, T. S., France, P., Fullagar, R., Giusti, D., Graham, S., Harris, M. D., Hawks, J., Health, S., Huffer, D., Kansa, E. C., Kansa, S. W., Madsen, M. E., Melcher, J., Negre, J., Neiman, F. D., Opitz, R., Orton, D. C., Przstupa, P., Raviele, M., Riel-Savatore, J., Riris, P., Romanowska, I., Smith, J., Strupler, N., Ullah, I. I., Van Vlack, H. G., VanValkenburgh, N., Watrall, E. C., Webster, C., Wells, J., Winters, J., Wren, C. D. 2017. Open Science in Archaeology. *SAArchaeological record* 17(4): 8-14.

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., Yarkoni, T. 2016. How open science helps researchers succeed. *eLife* 5: e16800.

Mehrer, M. W., Wescott, K. L. 2006. *GIS and archaeological site location modeling*. Boca Raton: CRC Press.

Menard, S. 2001. *Applied logistic regression analysis*. Thousand Oaks, CA: Sage Publications.

Méndez Fernández, F., González Méndez, M., Amado Reino, X. 1995. Control arqueológico del oleoducto Coruña-Vigo: fase I, trabajos previos y superficiales. *Actas del XXII Congreso Nacional de Arqueología (Vigo 1993)* 1: 293-296.

Méndez Fernández, F. 1998. Definición y análisis de poblados de la Edad del Bronce en Galicia, en R. Fábregas Valcarce (ed.) *A Idade do Bronce en Galicia: novas perspectivas*: 153-190. A Coruña: Seminario de Sargadelos.

Minetti, A., Moia, C., Roi, G., Susta, D., Ferreretti, G. 2002. Energy Cost of Walking and Running at Extreme Uphill and Downhill Slopes. *Journal of Applied Physiology* 93(3): 1039-1046.

Mlekuz, D. 2004. Listening to landscapes: modelling past soundscapes in GIS. *Internet Archaeology* 16. https://doi.org/10.11141/ia.16.6

Mlekuz, D. 2014. Exploring the topography of movement, en S. Polla y P. Verhagen (eds) *Computational Approaches to the Study of Movement in Archaeology: Theory, Practice and Interpretation of Factors and Effects of Long-Term Landscape Formation and Transformation*: 5-21. Berlin: De Gruyter.

Moldes Teo, F. J. 1995. *Tecnología de los sistemas de información geográfica*. Madrid: Ra-Ma.

Moreno Gallo, M. Á. 2004. *Megalitismo y geografía. Análisis de los factores de localización espacial de los dólmenes de la provincia de Burgos.* Valladolid: Universidad de Valladolid. Tesis doctoral inédita.

Murrieta Flores, P. 2012. Understanding human movement through spatial technologies: the role of natural areas of transit in the Late Prehistory of South-western Iberia. *Trabajos de Prehistoria* 69(1): 103-122.

Murrieta-Flores, P., García Sanjuán, L., Wheatley, D. 2014. Modelando la movilidad y el movimiento a través de los SIG. Estudios de caso en la Prehistoria Reciente de Sierra Morena Occidental, en E. García Alfonso (ed.) *Movilidad, Contacto Y Cambio. II Congreso de Prehistoria de Andalucía. Antequera, 15, 16 Y 17 de Febrero de 2012*: 77-92. Sevilla: Junta de Andalucía.

Naismith, W. W. 1982. Untitled. *Scottish Mountaineering Club Journal* 2: 135.

Nakoinz, O., Knitter, D. 2016. *Modelling Human Behaviour in Landscapes*. New York: Springer.

Negré, J., Muñoz, F., Barceló, J. A. 2018. A Cost-Based Ripley's K Function to Assess Social Strategies in Settlement Patterning. *Journal of Archaeological Method and Theory* 25: 777–794.

Novák, D. 2014. *Local Relief Model (LRM) Toolbox for ArcGIS*. https://onedrive.live.com/?cid=4a4965781c37ffac&id=4A4965781C37FFAC%2190606&authkey=%21AKcM9WetyjhaEAA.

Nüsslein, A., Nuninger, L., Verhagen, P. 2020. To Boldly Go Where No One Has Gone Before: Integrating Site Location Analysis and Predictive Modelling, the Hierarchical Types Map, en J. B. Glover, J. Moss y D. Rissolo (eds) *Digital Archaeologies, Material Worlds (Past and Present). Proceedings of the 45th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*: 15-31. Tübingen: Tübingen University Press.

O'Sullivan, D., Unwin, D. J. 2003. *Geographic Information Analysis*. New Jersey: John Wiley & Sons.

O'Sullivan, D., Perry, G. L. W. 2013. *Spatial Simulation: Exploring Pattern and Process*. Chichester: John Wiley & Sons.

Ogburn, D. E. 2006. Assessing the level of visibility of cultural objects in past landscapes. *Journal of Archaeological Science* 33(3): 405-413.

Olaya, V. 2014. *Sistemas de Información Geográfica*. https://volaya.github.io/libro-sig/

Openshaw, S. 1983. *The Modifiable Areal Unit Problem*. Norwich: Geobooks.

Orejas Saco del Valle, A. 1995. *Del marco geográfico a la Arqueología del paisaje: la aportación de la fotografía aérea*. Madrid: CSIC.

Orengo, H. A., Livarda, A. 2016. The seeds of commerce: a network analysis-based approach to the Romano-British transport system. *Journal of Archaeological Science* 66: 21-35.

Orengo, H. A., García-Molsosa, A. 2019. A brave new world for archaeological survey: Automated machine learning-based potsherd detection using high-resolution drone imagery. *Journal of Archaeological Science* 112: 105013.

Orton, C. 1988. *Matemáticas para arqueólogos*. Madrid: Alianza Editorial.

Orton, C. 2004. Point pattern analysis revisited. *Archeologia e Calcolatori* 15: 299-315.

Orton, C. 2006. Adding value to GIS - some spatial-analytical techniques and their applications, en *Reading historical spatial information from around the world: studies of culture and civilization based on geographic information systems data*: 3-16. Hong Kong: Research Center for Japanese Studies.

Paliou, E. 2013. Reconsidering the concept of visualscapes: recent advances in three-dimensional visibility analysis, en A. Bevan y M. Lake (eds) *Computational Approaches to Archaeological Spaces*: 243-263. Walnut Creek, California: Left Coast Press.

Palmisano, A. 2014. Zooming patterns among the scales: a statistics technique to detect spatial patterns among settlements, en G. Earl, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska, D. Wheatley (eds) *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*: 348-356. Southampton: Amsterdam University Press.

Parcak, S. 2009. *Satellite Remote Sensing for Archaeology*. London: Routledge.

Parcak, S. 2021. *La arqueología desde el espacio*. Barcelona: Ariel.

Parcero Oubiña, C. 2003. Looking forward in anger: social and political transformations in the Iron Age of the North-Western Iberian Peninsula. *European Journal of Archaeology* 6: 267–299.

Parcero Oubiña, C., Fábrega Álvarez, P. 2006. Diseño metodológico para el análisis locacional de asentamientos a través de un SIG de base raster, en I. Grau Mira (ed.) *La aplicación de los SIG en la Arqueología del Paisaje*: 69-91. Alicante: Servicio de Publicaciones Universidad de Alicante.

Parcero Oubiña, C., González Pérez, C. 2007. Los SIG y la gestión de la información arqueológica. Sistemas de información geográfica y análisis arqueológico del territorio, en V. Mayoral Herrera y S. Celestino Pérez (eds) *Tecnologías de Información geográfica y análisis arqueológico*

del territorio. Actas del V Simposio Internacional de Arqueología de Mérida: 481-489. http://hdl.handle.net/10261/140676

Parcero Oubiña, C., Armada, X. L., Ayán Vila, X. M. 2017. Castros en la escalera: el Noroeste entre la normalidad y la indiferencia, en S. Celestino Pérez (coord.) *La Protohistoria en la Península Ibérica*: 815-878. Madrid: Akal.

Parcero Oubiña, C., Nión Álvarez, S. 2021. Forms of settlement inequality over space. A GIS-based method for measuring differences among settlements. *Journal of Archaeological Science: Reports* 35: 102739.

Pardo-Gordó, S., Cortell-Nicolau, A. 2023. ¿Falta de concienciación o desconocimiento? La transparencia y la reproducibilidad en la Arqueología Computacional. *Vegueta: Anuario de la Facultad de Geografía e Historia* 23(1): 243-263.

Peebles, C. S. 1973. The sites and their setting, en C. S. Peebles (ed) *Moundville: the organisation of a prehistoric community and culture*. Ontario: University of Windsor Press.

Pélissier, R., Goreaud, F. 2001. A practical approach to the study of spatial structure in simple cases of heterogeneous vegetation. *Journal of Vegetation Science* 12: 99-108.

Peña Monne, J. L., Rodanés Vicente, J. M., Mazo Pérez, C., Montes Ramírez, M. L. 1986. La fotografía aérea vertical en blanco y negro y su aplicación a la prospección arqueológica y geoarqueológica, en A. Ubieto Arteta (coord.) *Metodología de la investigación científica sobre fuentes aragonesas: actas de las I Jornadas celebradas en Monzón, del 8 al 20 de diciembre de 1985*: 219-228. Zaragoza: Universidad de Zaragoza.

Pharr, M., Green, S. 2004. Ambient occlusion, en F. E. E. Randima (ed) *GPU Gems*: 279-292. New York: Addison Wesley.

Pickles, J. 1999. Arguments, debates, and dialogues: the GIS-social theory debate and the concern for alternatives, en P. Longley, M. Goodchild, D. J. Maguire y D. W. Rhind (eds) *Geographical Information Systems. Principles and Technical Issues*: 49-60. New York: John Wiley & Sons.

Polla, S., Verhagen, P. 2014. *Computational approaches to the study of movement in archaeology. Theory, practice and interpretation of factors and effects of long-term landscape formation and transformation*. Berlin: De Gruyter.

Premo, L. S. 2004. Local spatial autocorrelation statistics quantify multi-scale patterns in distributional data: an example from the Maya Lowlands. *Journal of Archaeological Science* 31: 855-866.

Rášová, A. 2014. Fuzzy viewshed, probable viewshed, and their use in the analysis of prehistoric monuments placement in Western Slovakia, en J. Huerta, S. Schade, C. Granel y C. Castellón (eds) *Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science*. Castellón, Springer.

Remondino, F., Campana, S. 2014. *3D Recording and Modelling in Archaeology and Cultural Heritage. Theory and best practices*. Oxford: BAR International Series.

Renfrew, C. 1976. Megaliths, territories and populations, en S. J. De Laet (ed) *Acculturation and continuity in Atlantic Europe mainly during the Neolithic period and the Bronze Age: papers presented at the IV Atlantic Colloquium. Ghent*: 198-220. Bruges: De Tempel.

Rennell, R. 2012. Landscape, experience and GIS: exploring the potential for methodological dialogue. *Journal of Archaeological Method and Theory* 19(4): 510-525.

Richards-Rissetto, H., Landau, K. 2014. Movement as a means of social (re) production: using GIS to measure social integration across urban landscapes. *Journal of Archaeological Science* 41: 365-375.

Ripley, B. D. 1976. The second-order analysis of stationary point processes. *Journal of Applied Probability* 13: 255-266.

Ripley, B. D. 1977. Modelling Spatial Patterns. *Journal of the Royal Statistical Society* 2: 172-212.

Ripley, B. D. 1981. *Spatial Statistics*. New York: John Willey and Sons.

Riris, P. 2020. Spatial structure among the geometric earthworks of western Amazonia (Acre, Brazil). *Journal of Anthropological Archaeology* 59: 101177.

Ritchie, S. 2020. *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science.* London: Bodley Head.

Rodríguez Casal, A. A., Eguileta Franco, J. M., Gómez Nistal, C., Ramos Alvite, E., Romaní Fariña, E. 1997. Metodología y primeras valoraciones de un proyecto interdisciplinar sobre el fenómeno tumular en la provincia de Lugo, en A. A. Rodríguez Casal (ed.) *O Neolítico Atlántico e as Orixes do Megalitismo. Actas do Coloquio Internacional (Santiago de Compostela, 1996)*: 521-536. Santiago de Compostela: Universidade de Santiago de Compostela.

Rodríguez del Cueto, F., Carrero Pazos, M. 2021. Límites y posibilidades de los análisis LiDAR aplicados al megalitismo asturiano. La revisión de cuatro conjuntos tumulares prehistóricos en el concejo de Salas (España). *Veleia* 38: 9-30.

Rodríguez Rellán, C., Fábregas Valcarce, R. 2015. Arte rupestre galaica: unha achega dende a estatística espacial e os SIX. *Sémata. Ciencias sociais e humanidades* 27: 323-348.

Rodríguez Rellán, C., Fábregas Valcarce, R. 2019. Monuments on the move. Assessing megaliths' interaction with the North-Western Iberian landscapes, en M. Hinz, J. Müller, M. Wunderlich (eds) *Megaliths, Societies and Landscapes: Early Monumentality and Social Differentiation in Neolithic Europe*: 621-639. Bonn: Habelt.

Rodríguez Rellán, C., Fábregas Valcarce, R. 2023. Assesing the perceptibility of prehistoric monuments on their landscape. An exploratory approach using agent-based modelling. *Vegueta: Anuario de la Facultad de Geografía e Historia* 23(1): 115-145.

Roselló Bordoy, G. 1970. Prospecciones arqueológicas en Mallorca mediante fotografía aérea. *Pyrenae* 6: 59-64.

Rubio-Campillo, X., Cela, J. M., Hernández Cardona, F. X. 2012. "Simulating" Archaeologists? Using Agent-based Modelling to Improve Battlefield Excavations. *Journal of Archaeological Science* 39: 347-356.

Rubio-Campillo, X. 2017. El papel de la simulación en la arqueología actual, en D. Jiménez-Badillo (ed.) *Arqueología computacional. Nuevos enfoques para la documentación, análisis y difusión del patrimonio cultural*: 51-58. México: Instituto Nacional de Antropología e Historia.

Ruiz Zapatero, G. 1988. La prospección arqueológica en España: Pasado, Presente y Futuro. *Arqueología Espacial* 12: 33-47.

Ruiz Zapatero, G., Burillo Mozota, F. 1988. Metodología para la investigación en Arqueología Territorial. *Munibe* 6: 45-64.

San José Albacete, A. 2011. *Procesamiento de datos LiDAR con ArcGIS Desktop 10*. Madrid: Universidad Complutense de Madrid.

Sánchez Palencia, F. J., Orejas Saco del Valle, A. 1991. *Fotointerpretación arqueológica: el estudio del territorio*. Madrid: Arqueología.

Sánchez Palencia, F. J., Fernández Posse, M. D. 1992. Fotointerpretación aplicada a la prospección arqueológica los castros de la Valderia y La Cabrera (León), en VV.AA. *Jornadas sobre teledetección y geofísica aplicadas a la arqueología, Madrid 7-10 de mayo de 1986, Mérida 1-3 de octubre de 1987*: 175-187. Madrid: Instituto de Conservación y Restauración de Bienes Culturales.

Santos Estévez, M. 2008. *Petroglifos y paisaje social en la prehistoria reciente del noroeste de la Península Ibérica*. TAPA: CSIC.

Saxe, A. A. 1970. *Social dimensions of mortuary practices*. Michigan: University of Michigan.

Sayer, D., Wienhold, M. 2012. A GIS-investigation of four early anglo-saxon cemeteries: Ripley's K-function analysis of spatial groupings amongst graves. *Social Science Computer Review* 31(1): 71-89.

Schiffer, M. B. 1976. *Behavioral archaeology*. New York: Academic Press.

Schmidt, S. C., Marwick, B. 2020. Tool-Driven Revolutions in Archaeological Science. *Journal of Computer Applications in Archaeology* 31(1): 18-32.

Shennan, S. 1988. *Quantifying archaeology*. Edinburgh: Edinburgh University Press.

Shennan, S., Downey, S. S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., Thomas, M. G. 2013. Regional Population Collapse Followed Initial Agriculture Booms in Mid-Holocene Europe. *Nature Communications 4*: 2486.

Silverman, B. D. 1986. *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Spencer, C., Bevan, A. 2018. Settlement location models, archaeological survey data and social change in Bronze Age Crete. *Journal of Anthropological Archaeology* 52: 71-86.

St. Joseph, J. K. S. 1966. *The Uses of Air Photography*. London: John Baker.

Steelman, K. L., Carrera Ramírez, F., Fábregas Valcarce, R., Guilderson, T., Rowe, M. W. 2005. Direct radiocarbon dating of megalithic paints from north-west Iberia. *Antiquity* 79: 379-389.

Stopher, P. R., Meyburg, A. H. 1979. *Survey sampling and multivariate analysis for social scientists and engineers*. Lexington: Lexington Books.

Štular, B., Kokalj, Ž., Oštir, K., Nuninger, L. 2012. Visualization of lidar-derived relief models for detection of archaeological features. *Journal of Archaeological Science* 39(11): 3354-3360.

Tarboton, D. G. 1997. A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models. *Water Resources Research* 33(2): 309-319.

Tarini, M., Cigoni, P., Montani, C. 2006. Ambient Occlusion and Edge Cueing to Enhance Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics* 15(5): 1237-1244.

Taylor, P. J. 1990. Editorial comment. GKS. *Political Geography Quarterly* 9: 211-212.

Thomas, J. 1993. The politics of vision and the archaeologies of landscape, en B. Bender (ed.) *Landscape: Politics and Perspectives*: 19-48. Oxford: Berg.

Thomas, J. 2004. *Archaeology and Modernity*. London: Routledge.

Thrane, H. 1998. The effects of the Bronze Age on the environment and culture in Scandinavia, en B. Hänsel (ed.) *Mensch und Umwelt in der Bronzezeit Europas*: 271-280. Kiel: Oetker-Voges.

Tilley, C. 2004. *The Materiality of Stone: Explorations in Landscape Phenomenology 1*. Oxford: Berg.

Tomlin, C. D. 1990. *Geographic Information Systems and Cartographic Modelling*. Englewood Cliffs: Prentice-Hall.

Tufte, E. R. 2001. *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.

Tzelepis, N., Loisios, D., Nakos, B. 2008. A methodology for creating analytical hill-shading by combining different lighting directions. *Proceedings of the 10th National Cartographic Conference, Hellenic Cartographic Society, Ioannina, 2008*: 1-10.

Vanzetti, A., Vidale, M., Gallinaro, M., Frayer, D.W., Bondioli, L. 2010. The iceman as a burial. *Antiquity* 84(325): 681-692.

Vaquero Lastres, J. 1991-1992. Del análisis del emplazamiento al estudio de la distribución de túmulos en el NW. *Brigantium* 7: 151-176.

Venables, W. N., Ripley, B. D. 2002. *Modern Applied Statistics with R*. New York: Springer.

Verhagen, P., Whitley, T. 2011. Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene. *Journal of Archaeological Method and Theory* 19(1): 1-52.

Verhagen, P. 2013. On the road to nowhere? Least cost paths, accessibility and the predictive modelling perspective, en F. Contreras, M. Farjas y F. J. Melero (eds) *Fusion of cultures.*

Proceedings of the 38th annual conference on computer applications and quantitative methods in archaeology (Granada, Spain, April 2010): 383-389. Oxford: Archaeopress.

Verhagen, P. 2018. Predictive Modeling, en S. L. López Varela (ed.) *The Encyclopedia of Archaeological Sciences*: 1-3. New Yersey: JohnWiley & Sons.

Verhoeven, G., Loenders, J. 2006. Looking through Black-Tinted Glasses-A Remotely Controlled Infrared Eye in the Sky, en S. Campana y M. Forte (eds) *From Space to Place: 2nd International Conference on Remote Sensing in Archaeology*: 73-79. Oxford: British Archaeological Reports.

Vicent García, J. M. 1997. The Island Filter Model Revisited, en M. S. Balmuth, A. Gilman y L. Prados-Torreira (eds) *Encounters and transformations. The Archaeology of Iberia in Transition*: 1-13. Sheffield: Sheffield Academic Press.

Vilas Estévez, B., Carrero Pazos, M., Eguileta Franco, J. M., Hast, A. 2015. Recreando con técnicas de fotogrametría histórica unha paisaxe megalítica asolagada: o caso do encoro do Salas (Ourense). *Férvedes* 8: 9-16.

Villoch Vázquez, V. 1995. Contribución al análisis del emplazamiento tumular: La necrópolis de As Travesas (Abegondo-A Coruña). *Minius* 4: 31-43.

Visentin, D., Carrer, F. 2017. Evaluating Mesolithic Settlement Patterns in Mountain Environments. *Archeologia e Calcolatori* 28: 129–145.

Vita Finzi, C., Higgs, E. 1970. Prehistoric Economy in the Mount Carmel area of Palestina: Site-Catchment Analysis. *Proceedings of the Prehistoric Society* 36: 1-37.

Warren, R. E., Asch, D. L. 2000. Site location in the Eastern Prairie Peninsula, en K. L. Wescott y R. J. Brandon (eds) *Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit*: 6-36. London: Taylor & Francis.

Weiss, A. 2000. Topographic position and landforms analysis. Poster Presentation. San Diego: ESRI User Conference.

Wescott, K. L., Brandon, R. J. 1999. *Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit*. London: Taylor & Francis.

Whallon, R. 1974. Spatial analysis of occupation floors. I. Applications of dimensional analysis of variance. *American Antiquity* 38: 320-328.

Wheatley, D. 1993. Going over old ground: GIS, archaeological theory and the act of perception, en J. Andersen, T. Madsen, I. Scollar (eds) *Computing the Past: Computer Applications and Quantitative Methods in Archaeology 1992*: 133-138. Aarhus: Aarhus University Press.

Wheatley, D. 1995. Cumulative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological application, en G. Lock y Z. Stančič (eds) *Archaeology and Geographical Information Systems: A European Perspective*: 171-186. London: Taylor & Francis.

Wheatley, D. 2000. Spatial technology and archaeological theory revisited, en K. Lockyear, J. T. Sly y V. Mihailescu-Birliba (eds) *CAA'96 - Computer Applications and Quantitative Methods in Archaeology*: 123-132. Oxford: Archaeopress.

Wheatley, D., Gillings, M. 2000. Vision, perception and GIS: developing enriched approaches to the study of archaeological visibility, en G. Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies*: 1-27. Amsterdam: IOS Press.

Wheatley, D., Gillings, M. 2002. *Spatial Technology and Archaeology. The archaeological applications of GIS*. London: Routledge

White, D. A., Barber, S. B. 2012. Geospatial modeling of pedestrian transportation networks: a case study from precolumbian Oaxaca, Mexico. *Journal of Archaeological Science* 39: 2684-2696.

Wickham, H., Grolemund, G. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Beijing: O'Reilly.

Wilke, C. O. 2019. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. Sebastopol: O'Reilly.

Wilkinson, M., Dumontier, M., Aalbersberg, I., *et al.* 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3: 160018.

Wilkinson, T. J. 2003. *Archaeological Landscapes of the Near East*. Tucson: University of Arizona Press.

Wilson, A. G. 2000. *Complex Spatial Systems*. London: Pearson Education.

Wilson, D. R. 2000. *Air photo interpretation for archaeologists*. Stroud: History Press.

Wilson, J. P., Gallant, J. C. 2000. *Terrain Analysis: Principles and Applications*. New York: John Wiley and Sons.

Wiseman, J. R. 2007. *Remote Sensing in Archaeology*. New York: Springer.

Woodman, P. E. 2000. A predictive model for Mesolithic site location on Islay using logistic regression and GIS, en S. J. E. Mithen (ed.) *Hunter-gatherer Landscape Archaeology: The southern Hebrides Mesolithic Project 1988-1998. vol. 2, Archaeological Fieldwork on Colonsay, Computer Modelling, Experimental Archaeology and Final Interpretations*: 445-464. Cambridge: McDonald Institute for Archaeological Research.

Xie, Y. 2016. *Bookdown. Authoring Books and Technical Documents with R Markdown*. Boca Raton: CRC Press.

Xie, Y., Allaire, J. J., Grolemund, G. 2018. *R Markdown: The Definitive Guide*. Boca Raton: CRC Press.

Yu, C., Lee, J., Munro-Stasiuk, M. 2003. Extensions to least-cost path algorithms for roadway planning. *International Journal of Geographical Information Science* 17: 361-376.

Yubero-Gómez, M., Rubio Campillo, X., López Chaero, F. J., Esteve Gràcia, X. 2015. Mapping changes in late prehistoric landscapes: a case study in the northeastern Iberian Peninsula. *Journal of Anthropological Archaeology* 40: 123-134.

Zakšek, K., Oštir, K., Kokalj, Ž. 2011. Sky-view factor as a relief visualization technique. *Remote Sensing* 3(2): 398-415.

Zakšek, K., Oštir, K., Pehani, P., Kokalj, Ž., Polert, E. 2012. Hill shading based on anisotropic diffuse illumination, en *Symposium GIS Ostrava 2012*: 1-10. Ostrava: Technical University of Ostrava.

Zamora Merchán, M. 2011. Improving Methods for Viewshed Studies in Archaeology: The Vertical Angle, en E. Jerem, F. Redo y V Szeverényi (eds) *On the Road to Reconstructing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 36th International Conference*: 614-622. Budapest: Archeaeolingua.

Zubrow, E. 2007. Remote sensing, fractals, and cultural landscapes: an ethnographic prolegomenon using U2 imagery, en J. Wiseman y F. El-Baz (eds) *Remote sensing in Archaeology*. New York: Springer.